The Thesis Committee for Bernard George David
Certifies that this is the approved version of the following thesis:

Examining "Choice": Identifying STEM Course-Offerings and Course-Taking Patterns in Charter and Non-Charter Public Schools through Social Network Analysis and Community Detection

APPROVED BY
SUPERVISING COMMITTEE:

Michael P. Marder, Supervisor

Jill A. Marshall

Examining “Choice”: Identifying STEM Course-Offerings and Course-Taking Patterns in Charter and Non-Charter Public Schools through Social Network Analysis and Community Detection

by

Bernard George David

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

August 2018

Acknowledgements

This work would not have been possible without support, guidance, and mentorship from a number of people. First and foremost, I thank Dr. Jill Marshall and Dr. Michael Marder for their feedback throughout this project and for supporting my graduate study in physics. They have each worked tirelessly on my behalf, and they have challenged me to become a better educator, physicist, and researcher. Dr. Marshall and Dr. Marder, I am incredibly grateful for all you have done.

I also thank Dr. María González-Howard for her enthusiastic support of my initial ideas for this project during her social network course. Without her course and feedback, I would not have been exposed to the social network analysis used in this project.

Thank you to the graduate students who have paved the way for me in Dr. Marder's "EduPhysics Task Force." Dr. Caitlin Hamrock, thank you for your patience in training me in the ERC and for introducing me to the data and software used in this work. Dr. Sarah Stephens, thank you for listening to and challenging my initial ideas for this project, forcing me to better articulate a rationale for pursuing this study. Dr. Matthew Guthrie, thank you for enthusiastically supporting earlier conceptions of this work and for helping me identify resources that I could use to realize my ideas.

This process would have been unimaginably more difficult without encouragement from my STEM Education cohort. Ashley, Kemper, Kerri, Gareth, Ryan, and Sneha, I am humbly appreciative of your support and friendship. Graduate study would certainly not be the same without you.

Risa, thank you for your unending support in a variety of ways: from proofreading early versions of this work to your steadfast faith and belief in my ability to complete this work when I doubted my own capacity to do so.

Last, I thank my family, whose love and encouragement carried me throughout this project. Mom and Dad, thank you for believing in me and for encouraging me during difficulty. Danny and Sam, thank you for keeping me rooted in reality. I cannot imagine a better pair of siblings.

Abstract

Examining “Choice”: Identifying STEM Course-Offerings and Course-Taking Patterns in Charter and Non-Charter Public Schools through Social Network Analysis and Community Detection

Bernard George David, M.A.
The University of Texas at Austin, 2018

Supervisor: Michael P. Marder

Public charter schools and other market-based reforms in education are heralded by proponents of school choice as efforts that empower families to make decisions about their students’ education that are specifically tailored to their students’ individual needs. Charter schools, in particular, have been positioned by political proponents as a key component of reform efforts striving to expand school choice for families. As proponents have argued, families dissatisfied with local non-charter schools deemed underperforming can elect to enroll their students in charter schools, which are purported to have the flexibility to experiment with novel, innovative instructional models outside the constraints of the traditional public education system. Given the political momentum supporting the expansion of charter schools in the United States, it is important to understand the programmatic differences between charter and non-charter schools. More specifically, characterizing programmatic differences between charter and non-charter schools will give researchers insight into whether or not students enrolling in different types of schools have expanded or limited course options. Toward that end, this works draws upon methods from physics—specifically community detection in network data—to: 1) explore differences in STEM course offering between public charter and non-charter secondary schools in Texas;

and 2) characterize the similarities and differences in student STEM course-taking patterns between Texas charter and non-charter secondary schools. As an extension of these broader research aims, this thesis seeks specifically to investigate how emergent tracks (as determined by prominent course-taking patterns within schools) within Texas charter and non-charter schools serve to either promote or constrain student access to STEM disciplines.

Table of Contents

List of Tables ix

List of Figures x

Chapter 1: Introduction 1

STEM Education 4

Sociophysics and Eduphysics 8

Chapter 2: Background 11

Charter School Potential: Market and Institutional Theories 12

Charter School Research: Student Outcomes and Market Pressures 14

Tracking and Ability Grouping 17

Social Network Analysis and Community Detection 27

Constructing Sociograms 28

Community Detection in Networks 29

Comparing Community Detection Algorithms 33

Network Visualization 35

Chapter 3: Methods 37

Texas Statewide STEM Course-Taking 40

STEM Course-Offerings in Charter and Non-Charter Schools 40

Student STEM Course-Taking in Charter and Non-Charter Schools 42

Chapter 4: Results 45

Texas Statewide STEM Course-Taking 45

STEM Course-Offerings in Charter and Non-Charter Schools 49

Student STEM Course-Taking in Charter and Non-Charter Schools 54

Chapter 5: Discussion 62

STEM Course-Offerings in Charter and Non-Charter Schools 62

STEM Course-Taking Patterns in Charter and Non-Charter Schools 64

Chapter 6: Conclusion 66
Limitations and Future Directions 67

Bibliography 68

List of Tables

Table 1. Proportion of bachelor's degrees conferred to students graduating from four-year post-secondary institutions in 2013-2014 by field and race/ethnicity. 5

Table 2. Average demographic characteristics of Texas charter and non-charter secondary schools. 38

Table 3. Courses associated with communities identified through school-level community detection. 52

Table 4. Coefficients for the multinomial regression model specified in Equation 16. 53

Table 5. Regression coefficients for the model specified in Equation 17. 54

Table 6. Average values of parameters used to cluster communities using the k-means algorithm. 56

Table 7. Average demographic characteristics of students associated with clusters of communities identified through k-means clustering by school sector. 57

Table 8. Coefficients from school-level multinomial logistic regression model using k-means clustering results as the outcome variable (Equation 18). . 58

Table 9. Coefficients for student level hierarchical logistic regression models specified by Equation 20. 59

Table 9 (continued). Coefficients for student level hierarchical logistic regression models specified by Equation 20. 60

List of Figures

Figure 1. Average science and math credits (in Carnegie Units) for high school graduates from select years between 1980 and 2010. 4

Figure 2. Average earnings in 2016 dollars for bachelor's degree holders by major. 6

Figure 3. Average earnings for bachelor's degree holders in all fields, STEM fields, and non-STEM fields. 7

Figure 4. Differences in advanced course-taking between Asian and all students in Texas public schools by percentage of economically disadvantaged students. 19

Figure 5. Difference in advanced course-taking between Black and all students in Texas public schools by the percentage of economically disadvantaged students. 20

Figure 6. Difference in advanced course-taking between Latinx and all students in Texas public schools by percentage of economically disadvantaged students. 21

Figure 7. Difference in advanced course-taking between White and all students in Texas public schools by percentage of economically disadvantaged students. 22

Figure 8. Difference in advanced course-taking between Asian and White students in Texas public schools by percentage of economically disadvantaged students. 23

Figure 9. Difference in advanced course-taking between Black and White students in Texas public schools by percentage of economically disadvantaged students. 24
Figure 10. Difference in advanced course-taking between Latinx and White students in Texas public schools by percentage of economically disadvantaged students. 25
Figure 11. A sociogram using the Fructherman-Reingold algorithm to place nodes. 36
Figure 12. Sociogram of STEM courses connected by all Texas public schools and colored by course category....................45
Figure 13. Sociogram of STEM courses connected by all Texas public schools and colored by multilevel community. 46
Figure 14. Sociogram of STEM courses connected by Texas charter schools and colored by course category....................47
Figure 15. Sociogram of STEM courses connected by Texas charter schools and colored by multilevel community. 47
Figure 16. Sociogram of STEM courses connected by Texas non-charter schools and colored by course category....................48
Figure 17. Sociogram of STEM courses connected by Texas non-charter schools and colored by multilevel community. 48
Figure 18. School-level sociogram in which edges are weighted by the number of courses shared by pairs of schools. The sociogram is colored by school sector....................50

Figure 19. School-level sociogram in which edges are weighted by the number of courses shared by pairs of schools. The sociogram is colored by community 51

Chapter 1: Introduction

Public charter schools emerged in the United States during the early 1990's. Throughout the decade both the number of charter schools and the number of students enrolling in charter schools steadily increased, a trend that continued throughout the 2000's and continues presently. According to the National Alliance for Public Charter Schools (NAPCS), over 3.2 million students were enrolled in roughly 7000 charter schools operating nationwide during the 2016-2017 school year (NAPCS, 2018). The prevalence of charter schools is unlikely to diminish anytime soon. At present, charter schools benefit from a political climate advocating for increased school choice and promoting charter schools as a key component of realizing this goal. A number of grant programs offered through the Department of Education's Office of Innovation and Improvement, created during Obama's presidency, are geared towards opening and expanding charter schools (Anderson, 2018) . More recently, Secretary of Education Betsy DeVos established grant funding guidelines aimed at expanding charter schools and school choice nationally, citing the need to provide families with alternatives to their neighborhood schools and empower families to enroll children in schools best suited to their students' needs (DeVos, 2017).

Charter schools are publicly-funded schools that typically, though not always, operate independently of local school districts. While public charter schools have greater autonomy than schools in local school districts, they operate under a contract with an authorizing organization that is responsible for holding charter schools accountable for student achievement. Proponents argue that independence from the traditional public-school system and autonomy over curriculum, financing, and staffing allow charter schools to innovate and develop novel educational models that promote student achievement more effectively than non-charter public schools[1] (Bierlein & Mulholland, 1994; Guggenheim,

[1] Research literature on charter schools often uses the term "traditional public schools" to differentiate between schools operating within public school districts and schools operating under a charter. This term suggests a homogeneity among the educational paradigms used in district schools, specifically an educational model that is standard and antiquated. In practice, there is tremendous variation between the educational models employed within district schools (as exemplified by the emergence of pilot and magnet

2010). This argument is founded upon market theory, which maintains that reducing restrictions within the public education system frees schools from stifling bureaucratic regulations and facilitates innovation through competition between schools (Berends & Donaldson, 2016; Lubienski, 2003). Proponents argue that with the freedom to innovate, charter schools have the potential to develop effective educational models that increase student learning, leading families to leave underperforming non-charter public schools and forcing these underperforming neighborhood schools to either improve or close.

A number of quantitative research studies have explored the impacts of charter schools on student outcomes, focusing on student achievement on standardized exams (Clark, Gleason, Tuttle, & Silverberg, 2015; Curto & Fryer, 2014; Gleason, Clark, Tuttle, & Dwoyer, 2010; Toma & Zimmer, 2012; Tuttle, Gleason, & Clark, 2012; Winters, 2012; Zimmer, Gill, Booker, Lavertu, & Witte, 2012) in addition to college enrollment and labor market outcomes (Dobbie & Fryer, 2016). These studies suggest that the impact of charter schools on student outcomes varies according to both the educational paradigms adopted by charter schools and the student populations served by charter schools. Several qualitative studies have explored how the introduction of market principles and competition to the public education sector has impacted the public education system in unintended ways (Jabbar, 2015, 2016; Lubienski, 2003; Winters, 2012). Several of these studies suggest that increases in test scores often attributed to charter schools may instead be artifacts of "cream-skimming," a practice in which charter schools recruit high-performing students from low-income backgrounds (Jabbar, 2015, 2016; Winters, 2015). Due to the fact that there is little consensus about the effects of charter schools on student outcomes, scholars advocate investigating the underlying conditions that may explain differences in student outcomes between charter and non-charter public schools (Berends, 2015; Berends & Donaldson, 2016).

In contrast to comparative studies exploring student achievement differences by school sector, Berends and Donaldson (2016) advocate for research to explore how school

---

schools within public school districts, for example). Given such variability, this work uses the term "non-charter public schools" to distinguish between district-operated schools and charter schools.

sector differences mediate observed student-outcome differences between charter and non-charter public schools. Toward that end, Berends and Donaldson (2016) explore how ability grouping in charter and non-charter schools influences student performance on standardized math exams. In their study, Berends and Donaldson (2016) infer ability groups by administering a survey to teachers in charter and non-charter schools in which teachers describe their instructional practices. Notably, these authors do not consider differences in course-offerings and student course-taking patterns between charter and non-charter public schools. This work seeks to extend the research conducted by Berends and Donaldson (2016) by looking at how students enroll in different sets of courses within charter and non-charter public schools in addition to how differences in course offerings by school sector mediate student course-taking patterns. Specifically, this work addresses the following research questions:

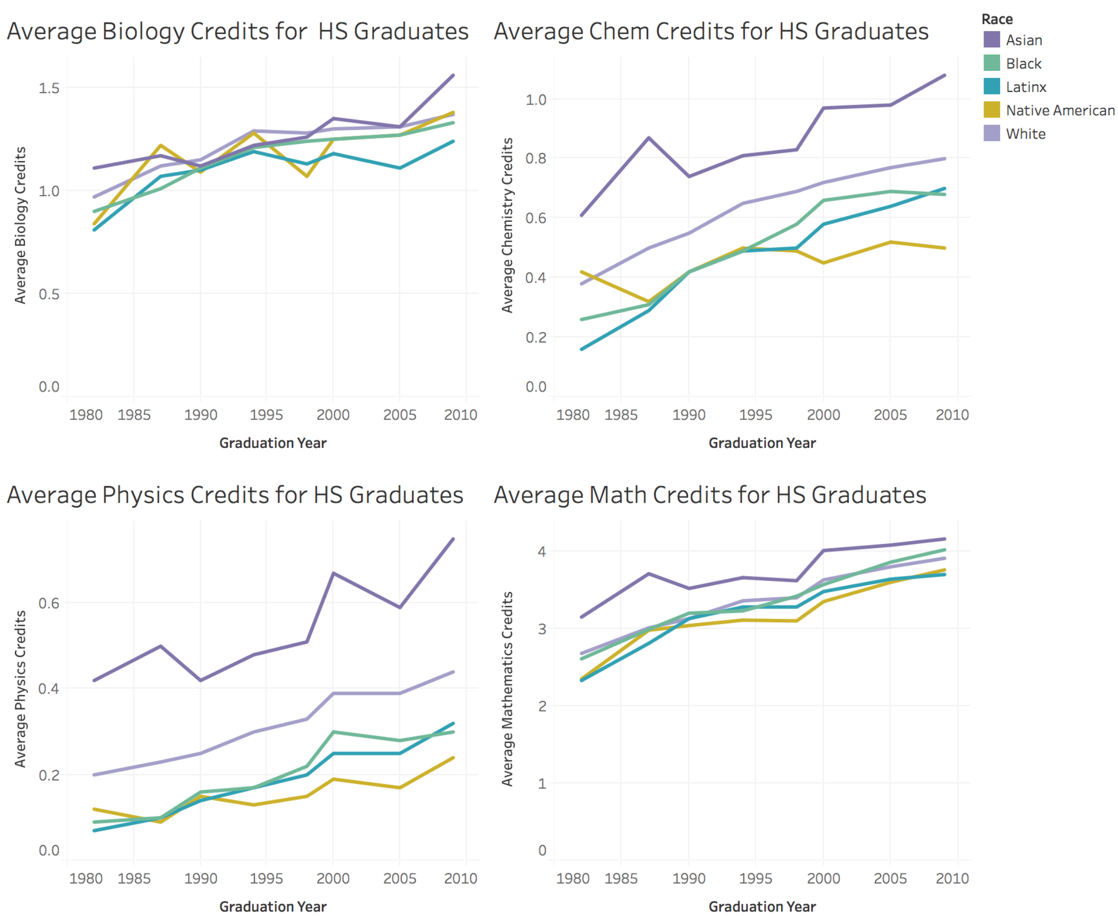
1. What are the programmatic differences between Texas charter and non-charter public schools, specifically in STEM (science, technology, engineering, and mathematics) course offerings?
2. What are the differences in STEM course-taking patterns between students enrolled in charter schools and students enrolled in non-charter public schools?

This work employs concepts and mathematical methods from physics in order to address these research questions. Specifically, this work combines social network analysis and community detection to analyze administrative educational data in Texas in an effort to identify STEM course-offerings and prominent student course-taking patterns in charter and non-charter public schools. A number of the algorithms considered in this work draw inspiration from physics, such as the use of resistors and spin glass systems to identify community structure in networks. Results from social network analysis and community detection are subsequently used as outcome variables in hierarchical linear models and multinomial logistic regression to ascertain the degree to which course-offerings and student course-taking patterns differ by school sector. Results indicate that charter schools are associated with an increase in the probability of a student taking advanced STEM course-sequences relative to course sequences defined as "college preparatory." In addition, charter schools are associated with an increased likelihood of students enrolling in course sequences that are more basic than college preparatory, in which students neither take advanced nor elective STEM courses. Finally, course-sequences in charter schools are more likely than non-charter schools to include exit from that school (specifically dropping out) and transfer to and from that school.

STEM EDUCATION

Criticism over the state of STEM Education often expressed in political discourse is based on the premise that the United States is underpreparing students for careers in STEM fields, which will ultimately result in the United States losing its status as a leader

Figure 1. Average science and math credits (in Carnegie Units) for high school graduates from select years between 1980 and 2010.

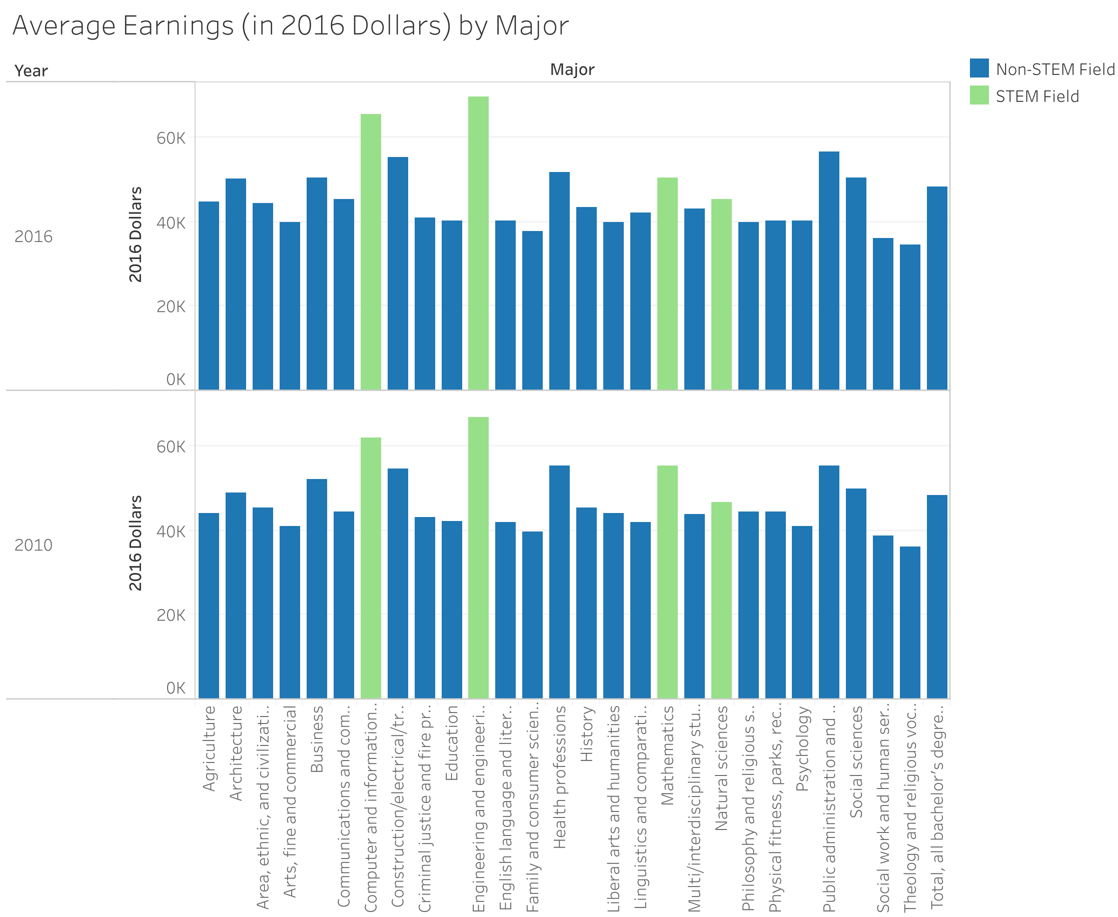


in scientific, engineering, and technological innovation (Feder, 2012; Office of the Press Secretary, 2009). United States students' subpar performances on the National Assessment of Educational Progress (NAEP) and Program for International Student Assessment (PISA) seemingly support this narrative and are often cited as evidence for a STEM shortage in the United States (National Science Board, 2016). Probing more deeply, however, evidence to support the existence of a STEM shortage in the United States is insufficient, and the United States' mediocre performance on NAEP and PISA is instead emblematic of discrepant engagement with the STEM disciplines according to racial and socioeconomic factors (Salzman, 2013; Teitelbaum, 2014). Such discrepant participation in STEM fields is evident as early as high school. Figure 1 displays the average number of STEM credits (in Carnegie Units) earned by high school graduates between 1982 and 2009 disaggregated by race (National Center for Education Statistics, 2018). On average, Asian and White students take more credits in the physical sciences (physics and chemistry) during high school than Black, Latinx, and Native American students. Differences in biology and math credits, however, are not as substantial.

Table 1. Proportion of bachelor's degrees conferred to students graduating from four-year post-secondary institutions in 2013-2014 by field and race/ethnicity.

	Native Am. & Alaska Native	Asian	Black	Latinx	Pacific Islander	White
All fields	*0.006*	*0.067*	*0.102*	*0.108*	*0.003*	*0.652*
Biological and biomedical sciences	0.004	0.157	0.074	0.096	0.003	0.610
Computer and information sciences	0.005	0.104	0.107	0.092	0.003	0.611
Engineering	0.003	0.11	0.040	0.088	0.002	0.646
Engineering technologies and engineering related fields	0.009	0.04	0.105	0.088	0.002	0.697
Health professions and related programs	0.005	0.074	0.119	0.089	0.004	0.676
Mathematics and statistics	0.003	0.103	0.048	0.083	0.001	0.620
Physical sciences and science technologies	0.006	0.101	0.053	0.076	0.002	0.686

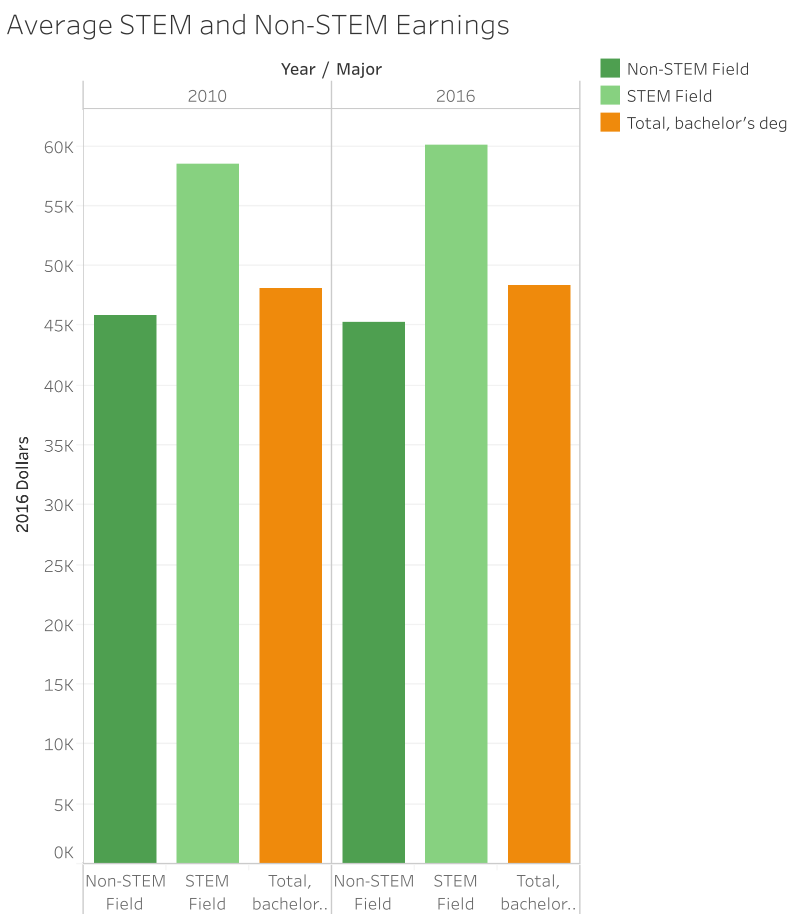
Figure 2. Average earnings in 2016 dollars for bachelor's degree holders by major.



Differences in STEM participation at the post-secondary level also exists along ethnic and socioeconomic lines. Table 1 displays the proportion of bachelor's degrees conferred by United States post-secondary institutions in 2013-2014 by race/ethnicity. Of the post-secondary student population, Black, Latinx, Pacific Islander, and American Indian/Alaska Native students are underrepresented in STEM disciplines—particularly in engineering, mathematics, and the physical sciences—while White and Asian students are overrepresented in these disciplines.

While scholars problematize rhetoric used to advance an agenda seeking to bolster preparation in STEM fields and fix the "leaky STEM pipeline," in which students—particularly women and underrepresented minority populations—discontinue their pursuits of STEM disciplines at educational junctures (Metcalf, 2010; Teitelbaum, 2014), inequitable access to and participation in STEM fields is cause for concern. Typically, STEM workforce and STEM pipeline literature defines STEM careers as those requiring post-secondary preparation in STEM, overlooking other careers that require STEM proficiency without a bachelor's degree. Rothwell (2013) reports that, as of 2011, half of all careers requiring STEM knowledge did not require workers to have a bachelor's degree in a STEM discipline. Moreover, these jobs pay average salaries that are 10% higher than

Figure 3. Average earnings for bachelor's degree holders in all fields, STEM fields, and non-STEM fields.



careers in other fields with equivalent educational requirements. Inequitable access to and participation in STEM can therefore inhibit students' ability to pursue jobs and careers that are more lucrative.

Differences in compensation between STEM and non-STEM fields are not limited to those that require less than a bachelor's degree, as Rothwell (2013) reports. Figure 2 and Figure 3 show the average earnings for bachelor's degree holders in STEM and non-STEM fields. Of note, STEM degrees are associated with higher average earnings than non-STEM fields (National Center for Education Statistics, 2018). Carnevale, Cheah, and Hanson (2015) also report that post-secondary degrees in STEM fields are associated with wages higher than the average for all fields. In addition to degrees in STEM fields, business and health degrees are also associated with higher than average wages (Carnevale et al., 2015).

Given the economic benefits of a career in STEM—both with and without post-secondary education in STEM fields—it is important to explore how discrepant participation in STEM fields by socioeconomic and ethnic boundaries serves to limit access to these economic benefits. Charter schools are positioned by proponents as an educational reform that can serve to promote social mobility; however, there is little research exploring the programmatic mechanisms by which this can be achieved (Berends, 2015; Berends & Donaldson, 2016; Berends, Goldring, Stein, & Cravens, 2010). Adequate preparation at STEM, at the secondary and post-secondary levels, is a viable way to promote social mobility, seeing as STEM careers are associated with higher than average wages. Understanding differences in course-offerings and course-taking patterns between charter and non-charter schools is important, as student access to advanced STEM coursework may lead to increased college enrollment and improved labor market outcomes.

SOCIOPHYSICS AND EDUPHYSICS

Schweitzer (2018) writes, "Generic modeling approaches that replicate physics insights, such as phase transitions and scaling laws, may reveal a lot about statistical physics but little about social dynamics. Merely using physical metaphors and analogies does not make physics applicable" (p. 41). However, quantitative frameworks from physics are instructive in exploring social phenomena, particularly as big data continues to become increasingly prevalent in modern society (Schweitzer, 2018). The application of physics to social phenomena—sociophysics—is particularly useful in analyzing and developing models from big data in addition to modeling complex networks.

In a similar vein, eduphysics utilizes theoretical principles in addition to computational and quantitative methodologies from physics to explore trends in educational data (Guthrie, 2018). Specifically, principles from statistical and fluid mechanics have been used to visualize longitudinal educational data and use students' test score flows to both evaluate the impacts of educational policy and predict long-term student achievement (Bendinelli & Marder, 2012; Marder & Bansal, 2009). Moreover, the use of clustering algorithms and machine learning have been used to classify public schools based upon their similarity in such a way that educational researchers can more meaningfully identify comparison groups when conducting statistical analyses on public schools (Guthrie, 2018).

One area in particular in which physics principles have been useful in developing analytical tools for examining social and educational phenomena is network analysis. Networks consist of agents, represented as nodes or vertices, whose interactions are represented by edges between them. The strength and direction of agents' interactions can also be modeled in network analysis. Beyond just modeling a network of agents at a macroscopic level, it is also instructive to uncover the underlying structures and substructures of the network. Toward this end, a number of physicists have developed and tested community detection algorithms in order to provide researchers with analytical tools to more deeply understand a network of interest (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008; Bruna & Li, 2017; Clauset, Newman, & Moore, 2004; Danon, Díaz-Guilera, Duch, & Arenas, 2005; Fortunato, 2010; Girvan & Newman, 2002; Lancichinetti & Fortunato, 2009; Lancichinetti, Fortunato, & Radicchi, 2008; Lancichinetti, Radicchi, Ramasco, & Fortunato, 2011; Newman, 2006; Newman & Girvan, 2004; Yang, Algesheimer, & Tessone, 2016). By constructing networks in which students (and schools) are connected to one another according to the number of courses they share in common, this work seeks to use community detection in an effort to begin to: 1) uncover how course-offerings inhibit or promote student engagement in STEM; and 2) investigate differences in course-taking patterns by school sector. Network analysis and the community detection algorithms developed using physical principles are well-suited for these goals.

Chapter 2: Background

Ample research exists investigating multiple facets of charter schools. A number of quantitative studies compare student outcomes between charter and non-charter schools, looking in particular at differences in student test score gains, college enrollment, and labor market outcomes by school sector (Clark et al., 2015; Dobbie & Fryer, 2016; Gleason et al., 2010; Tuttle et al., 2012). Qualitative research explores how the introduction of market pressures into the public education system has impacted the educational system writ large and influenced school leaders' decision-making strategies (Jabbar, 2016; Lubienski, 2003; Mungal, 2016). In addition, other studies have investigated whether or not the instructional conditions in charter schools support students' cognitive development (Finn et al., 2014; West, Gabrieli, Finn, Kraft, & Gabrieli, 2014) and cater to students' diverse sociocultural needs (Doyle & Feldman, 2006; McDermott & Nygreen, 2013; Modica, 2015).

Berends and Donaldson (2016) argue for researchers to move beyond comparing student outcomes by sector and to investigate the programmatic conditions within charter and non-charter schools that may be responsible for observed differences in student outcomes between them. Specifically, Berends and Donaldson (2016) explore how ability grouping in charter and non-charter schools mediate observed differences in students' math score gains. With the exception of Berends and Donaldson (2016), literature exploring tracking and ability grouping in schools is distinct from research on charter schools. This thesis seeks to expand upon the work of Berends and Donaldson (2016) by investigating STEM course-offerings and course-taking patterns in Texas charter and non-charter schools. As charter schools are positioned by proponents as schools of choice, a goal of this work is to elucidate how choice manifests for students by cataloguing how sector differences in course-offerings constrain or promote students' STEM course-taking options.

To evaluate course-offerings and student course-taking patterns in Texas charter and non-charter schools, this work analyzes administrative data collected by the Texas Education Agency (TEA). Social network analysis and community detection offer novel ways of identifying course-offerings and course-taking patterns in public schools and allow for comparisons to be made between school sector. In considering course-taking pattern differences between charter and non-charter schools, however, this work does not look specifically at either "tracking" or "ability grouping," as the data analyzed herein cannot be used to discern how or why students enrolled in certain STEM courses. Rather than attempting to ascertain how students are assigned to courses, this work identifies course-offerings and course-taking patterns in Texas charter and non-charter schools in order to offer insight into how programmatic differences by school sector potentially influence student outcomes.

This literature review is organized as follows: first market and institutional theories, which lead to different conclusions regarding the potential of charter schools to transform the public education system, are reviewed; next, research on charter school impacts is considered; third, tracking and ability grouping are discussed; and finally, social network analysis and community detection methods are introduced.

CHARTER SCHOOL POTENTIAL: MARKET AND INSTITUTIONAL THEORIES

Charter school proponents often rely on market theory to justify calls to create and expand charter schools (Berends et al., 2010; Bierlein & Mulholland, 1994; Henig, 1995; Lubienski, 2003). In market theory, consumers are free to choose goods and services based on price and quality. If a purveyor of goods and services overcharges for goods and services, or alternatively, if a purveyor's goods and services are of poor quality, consumers can turn elsewhere. In addition, entrepreneurs have the ability to invent products and offer services for lower costs and of greater quality should there be insufficient options currently available in the market. In this way, the consumer's ability to choose where to purchase goods and services incentivizes competition between providers who need to offer quality goods and services at reasonable prices or risk losing business.

As Lubienski (2003) describes, critics of the public education system argue that public education is a monopoly that stifles the abilities of individual schools to innovate. School choice proponents argue that the introduction of market-like competition to public education will serve to liberate schools from such limiting bureaucratic regulations, giving schools greater autonomy with which to innovate and create novel instructional paradigms that serve to meet students' needs with greater quality and efficiency (Friedman, 2009; Henig, 1995). With increased choice in the market place, families are able to enroll their students in schools that best serve their children's needs, and market pressures will force schools that are unable to innovate or meet the needs of students to close.

Skeptics contend, however, that public education is a "public good" and may not respond to market pressures in the way some economists suggest. Henig (1995) notes that education benefits the broader community by creating a better trained work force and a populace that is both prepared for civic engagement and able to lead fulfilling lives . These public benefits are not necessarily considered when families are left to invest in education as they see fit, and as a result, families may underinvest in their children's education resulting in decreased efficiency in public education. In addition, Lubienski (2003) notes that research on charter schools indicates that the innovation expected due to market pressures is often limited to schools' organizational structure, but is not evident in novel curricula or instruction. Lubienski (2003) argues that "curricular conformity and instructional standardization may in fact be caused by the very market mechanisms that were unleashed to address those ills" (p. 397).

That charter schools revert to established curricular and instructional norms is predicted by institutional theory, in which an organization's "legitimacy is derived from conformity to the normatively held rules and scripts of the institutional environment, rather than instructional effectiveness" (Huerta & Zuckerman, 2009, p. 414). According to institutional theory, competition and innovation introduced to the public education system through market pressures are not as powerful as charter school advocates suggest. The bureaucracy of the public education system has long established normative schooling, and schools, regardless of sector, adhere to practices that serve to legitimize these organizations as educational entities (Berends & Donaldson, 2016; Huerta & Zuckerman, 2009). While charter schools have greater autonomy with which they have the potential to innovate, established norms and state regulations within education may prove insurmountable.

CHARTER SCHOOL RESEARCH: STUDENT OUTCOMES AND MARKET PRESSURES

Literature exploring differences in student achievement between charter and non-charter public schools suggests the impact of charter schools upon student outcomes is contextual at best (Gleason et al., 2010; Zimmer et al., 2012). In a study looking at student achievement in charter and non-charter schools across seven states, Zimmer et al. (2012) conclude there is little or no difference by sector, as charter schools in some states tend to increase students' test scores in math and reading, while charter schools in other states yield decreased students' standardized test scores. Gleason et al. (2010) used a natural experiment including 36 charter middle schools from 15 states to compare outcomes between students selected to attend charter schools through a randomized lottery to students who were not selected from the lottery. On average, Gleason et al. (2010) found no differences in student achievement in math or reading between charter and non-charter public schools; however, differences in student achievement were found when disaggregating schools by student populations. Specifically, charter schools serving underperforming, low-income students yielded positive student outcomes in mathematics, whereas charter schools serving high-performing, high-income students yielded negative student outcomes in mathematics and reading. In a separate article using data from the same study, Clark et al. (2015) describe that charter schools in urban settings improve student mathematics scores while charter schools in non-urban settings do not.

The ability of charter schools to improve student outcomes has also been shown to vary according to the educational paradigm adopted by the school, and charter schools have been created with a range of goals, from "rescue and recovery" schools serving students at risk for dropping out to college preparatory schools. Dobbie & Fryer (2016) examine a variety of outcomes—student test score gains, college enrollment, and early market labor outcomes—of charter school graduates. Specifically, "no excuses" charter schools—schools with longer school days, rigorous test preparation, and high behavioral standards—were found to increase student test score gains and college enrollment but have little noticeable impact upon graduates' future earnings. Other charter schools, by contrast, were associated with decreases in student test scores, college enrollment, and future earnings.

Additional evidence for the importance of a charter school's educational model on student achievement comes from Curto and Fryer (2014) who studied student math and reading achievement at SEED, an urban, boarding, college-preparatory charter school in Washington, D.C[2]. When comparing students randomly selected for attendance through lottery to students not selected for attendance, Curto and Fryer (2014) report students attending SEED had increases in both mathematics and reading scores.

In comparing how two governing organizations in New Orleans, in which a majority of public schools are charter schools, regulate market-based competition, Jabbar (2016) concludes that family choice in a "deregulated" educational system is heavily moderated by the structures and policies set forth by organizations governing schools. School leaders' perception of the competitive landscape is shaped by the policies established by governing bodies, and thus influences how leaders respond to these pressures. In a separate article, Jabbar (2015) explores how school leaders in New Orleans respond to competition. Jabbar (2015) finds that leaders feel pressure to recruit and retain students in order to receive funding for these students. In order to recruit students, leaders adopt a number of strategies: 1) improving test scores through changes to the academic program; 2) catering academic programming to niche interests; 3) employing marketing campaigns to recruit students; and 4) cream-skimming, a practice in which schools target recruiting efforts towards high-achieving students from low-income backgrounds. That some charter schools market toward high-achieving students, or alternatively do not market toward low-achieving students in low-income neighborhoods, both exacerbates inequities in access to education and raises concerns regarding the true nature of charter school impacts on student outcomes (Jabbar, 2016; Lacireno-Paquet, Holyoke, Moser, & Henig, 2002). Rather than inducing innovation within the public-school sector, market pressures may instead encourage schools to engage in recruiting practices that serve to give them a competitive advantage by ignoring students who are deemed underachieving.

[2] Before enrolling in graduate school, I taught physical science and physics to 8th and 9th grade students, respectively, at SEED. The study by Curto and Fryer (2014) was conducted prior to my employment at SEED.

In addition to concerns over the practice of "cream-skimming" at charter schools and the potential of this to account for positive impacts on student achievement attributed to charter schools, other scholars express concern that charter schools may achieve increases in student test scores by "teaching to the test" (Finn et al., 2014; West et al., 2014). Finn et al. (2014) compared student gains on cognitive measures between charter and non-charter schools, observing that although charter schools increase students' standardized test scores, there were no differences in cognitive growth between charter and non-charter school students. Such a finding suggests that charter schools achieve test score gains without enhancing students' cognitive skills.

Some scholars argue that charter schools that increase college enrollment do so through "new paternalistic" approaches to education in which low-income students of color are taught how to exhibit middle-class values through intensive character education that supplements academic programming (Curto & Fryer, 2014; McDermott & Nygreen, 2013). While new paternalistic approaches to urban education have been positively received due to their purported ability to promote social mobility by enhancing students' social capital, McDermott and Nygreen (2013) note that these practices suppress low-income students' capacity to develop self-advocacy within the educational system, a critical aspect of what helps middle class children successfully navigate schooling. Modica (2015) also observed that instructional conditions in a diverse charter school pressured students of color to act "White" in order to be perceived as academically capable.

Berends and Donaldson (2016) underscore the importance of moving beyond comparing student outcomes between charter and non-charter public schools and instead focusing research efforts on developing a better understanding of the differences between charter and non-charter public schools that may be responsible for observed differences in student achievement. Berends and Donaldson (2016) write, "placing charter schools in a horserace between sectors (i.e., comparing charter schools to traditional public schools) is not as helpful as understanding the conditions under which school effects—traditional, charter, private, etc.—occur" (p. 3).

Toward this end, Berends and Donaldson (2016) explore how differences in ability grouping between charter and non-charter public schools explain differences in student achievement. With respect to differences in ability grouping between charter and non-charter public schools, Berends and Donaldson (2016) find that students in charter schools are more likely than students in non-charter public schools to be in a high ability group and less likely to be in an average ability group. While there are differences between mathematics achievement gains between students in high and low ability groups, this relationship does not differ between charter and non-charter public schools, suggesting that ability grouping practices between the two sectors are more alike than they are different, despite the different proportion of students enrolled in ability groups by sector (Berends & Donaldson, 2016).

TRACKING AND ABILITY GROUPING

Although often used interchangeably, tracking is the practice of sorting students into entire course sequences based upon their perceived academic aptitudes or prior achievement, whereas ability grouping refers to the practice of enrolling "students into classes on a subject-by-subject basis" based on their perceived ability (Berends & Donaldson, 2016, p. 7). Although the explicit practice of tracking has diminished in the United States, the vertical sequence of courses at the high school level has allowed ability grouping to continue (Heck, Price, & Thomas, 2004). Specifically, students who are deemed prepared may be able to enroll in more advanced coursework, while other students are forced to enroll in lower-level coursework. There is concern that ability grouping at the secondary level allows for inequitable access both to advanced coursework and post-secondary options. As Friedkin and Thomas (1997) describe, "evidence suggests that the odds of entering particular tracks are affected by many conditions, including the academic ability and socioeconomic origins of the students, and that tracks affect students' educational attainments, attitudes, decisions, and peer relations" (p. 239).

Using social network analysis, Heck et al. (2004) followed a cohort of 274 students throughout their high school careers and linked students to the courses they took each semester. In doing so, Heck et al. (2004) identified emergent course-taking patterns and found that enrollment in these pathways is both predicted by socioeconomic status and predictive of post-secondary plans. Specifically, advantaged students with high prior achievement are more likely to be enrolled in advanced course sequences, to have high college admissions exam scores, and to plan on attending colleges after graduating from high school.

McFarland (2006) similarly borrows from social network analysis in order to identify curricular trajectories in two high schools and explore mobility between these academic tracks. Notably, McFarland (2006) reports that the types of curricular trajectories can vary widely between schools as a function of what course offerings are available at that school. Moreover, the structures of course sequences play an important role in determining whether or not students are able to move between and among different tracks within a school. For example, students in a school with increased differentiation in higher-level courses have greater opportunities to move to advanced tracks, whereas students in schools with limited higher-level courses must compete for spots in these courses.

Differential access to advanced courses in the public-school system can easily be seen using publicly available educational data. Plots generated using educational data from Texas show clear racial/ethnic differences in the rates at which students are enrolled in advanced coursework, corroborating research that suggests advantaged students with high socioeconomic status are more likely to enroll in advanced course sequences (Friedkin & Thomas, 1997; Heck et al., 2004).

Figure 4 plots the difference between the percentages of Asian students and all students enrolling in advanced coursework within a given school by the percentage of students who students who are economically disadvantaged within the school, as identified by the number of students qualifying for free and reduced lunch (FRL). The plot is colored by the percentage of underrepresented minority students, defined as the students who do not identify as White, enrolled at that school, and the size of the data points on the plot are proportional to each school's student population. In the plot, square data points represent charter schools and circular data points represent non-charter public schools.

Figure 4. Differences in advanced course-taking between Asian and all students in Texas public schools by percentage of economically disadvantaged students.

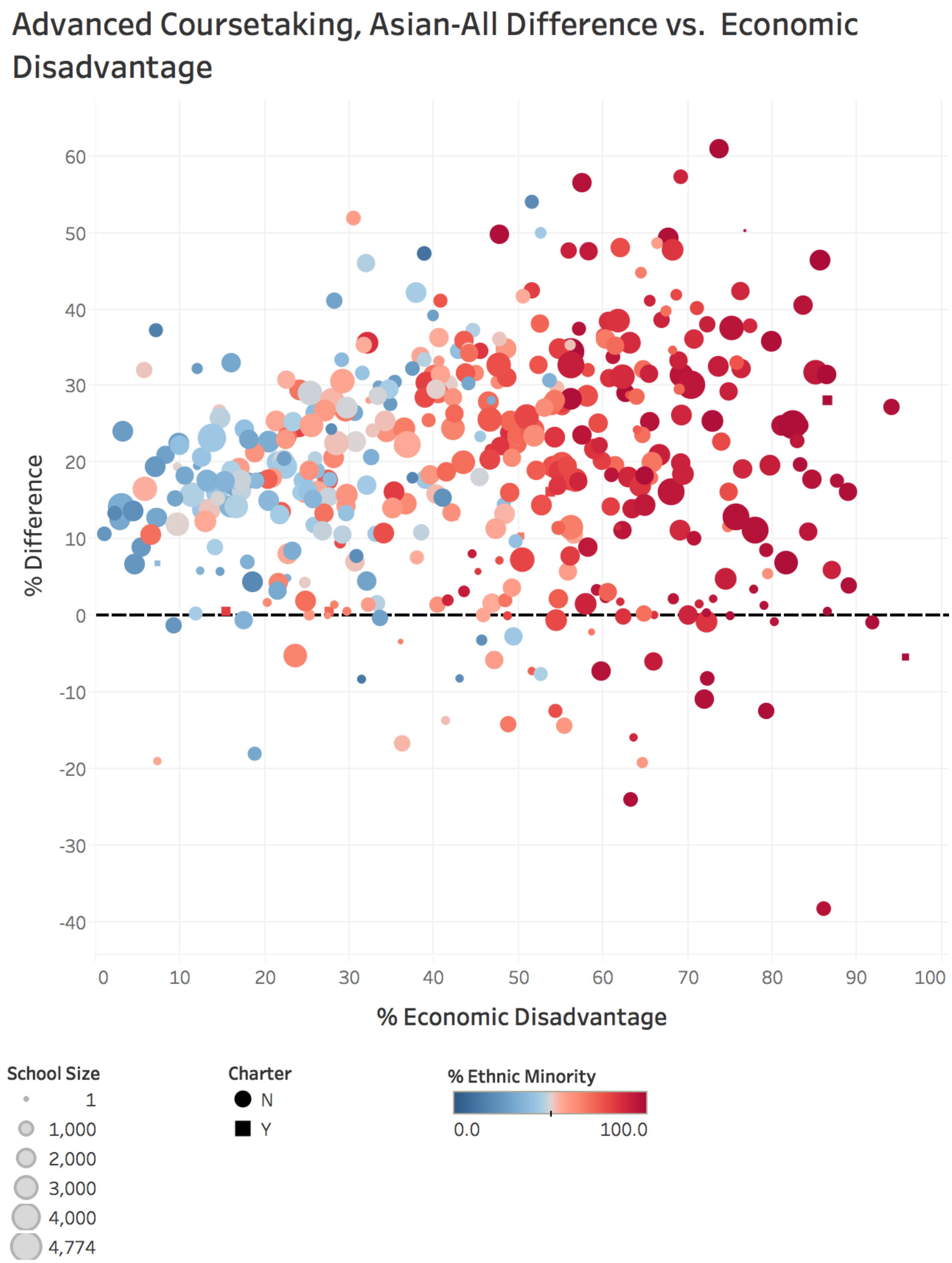


Figure 5. Difference in advanced course-taking between Black and all students in Texas public schools by the percentage of economically disadvantaged students.

Figure 6. Difference in advanced course-taking between Latinx and all students in Texas public schools by percentage of economically disadvantaged students.

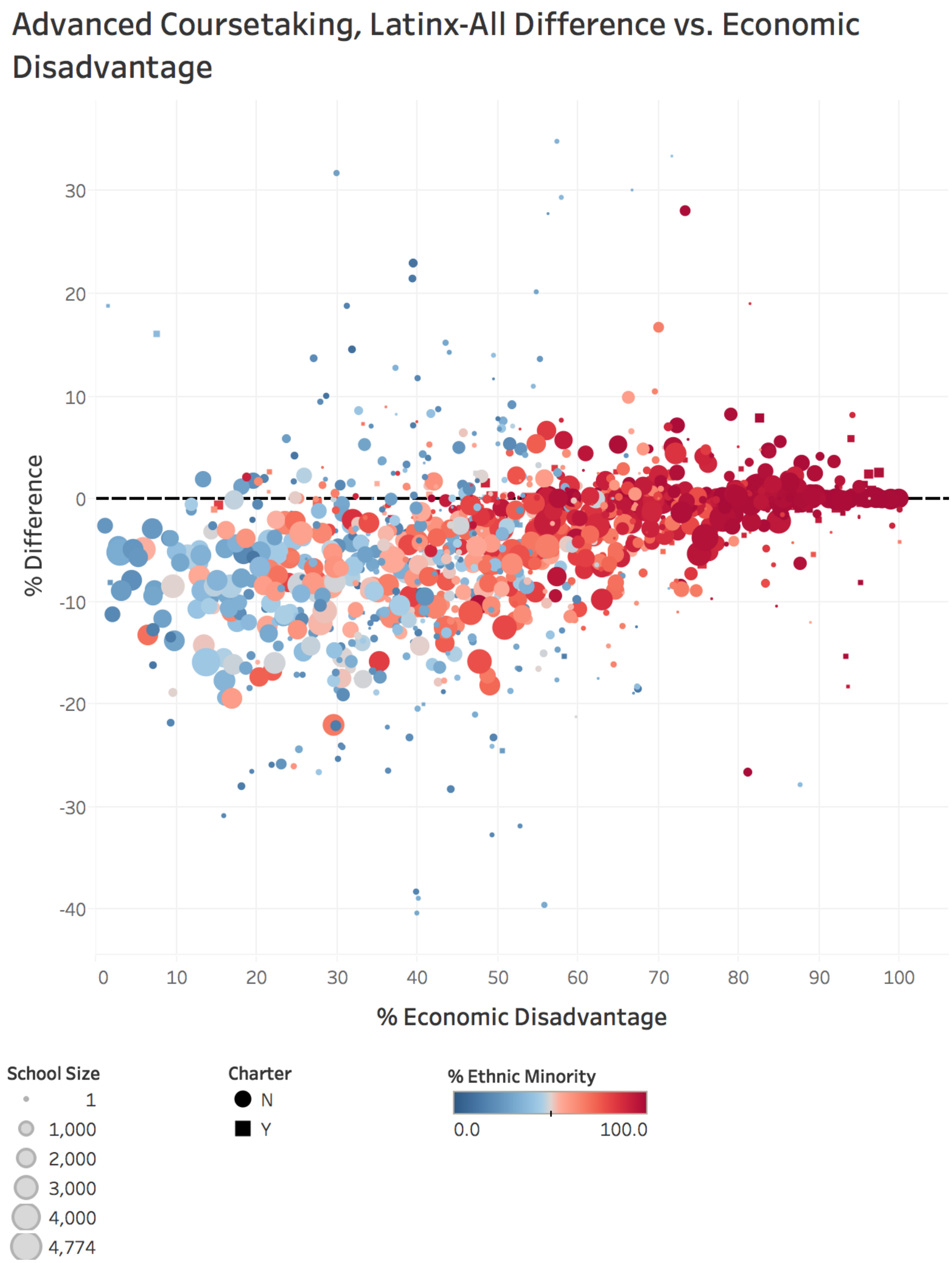


Figure 7. Difference in advanced course-taking between White and all students in Texas public schools by percentage of economically disadvantaged students.

Figure 8. Difference in advanced course-taking between Asian and White students in Texas public schools by percentage of economically disadvantaged students.

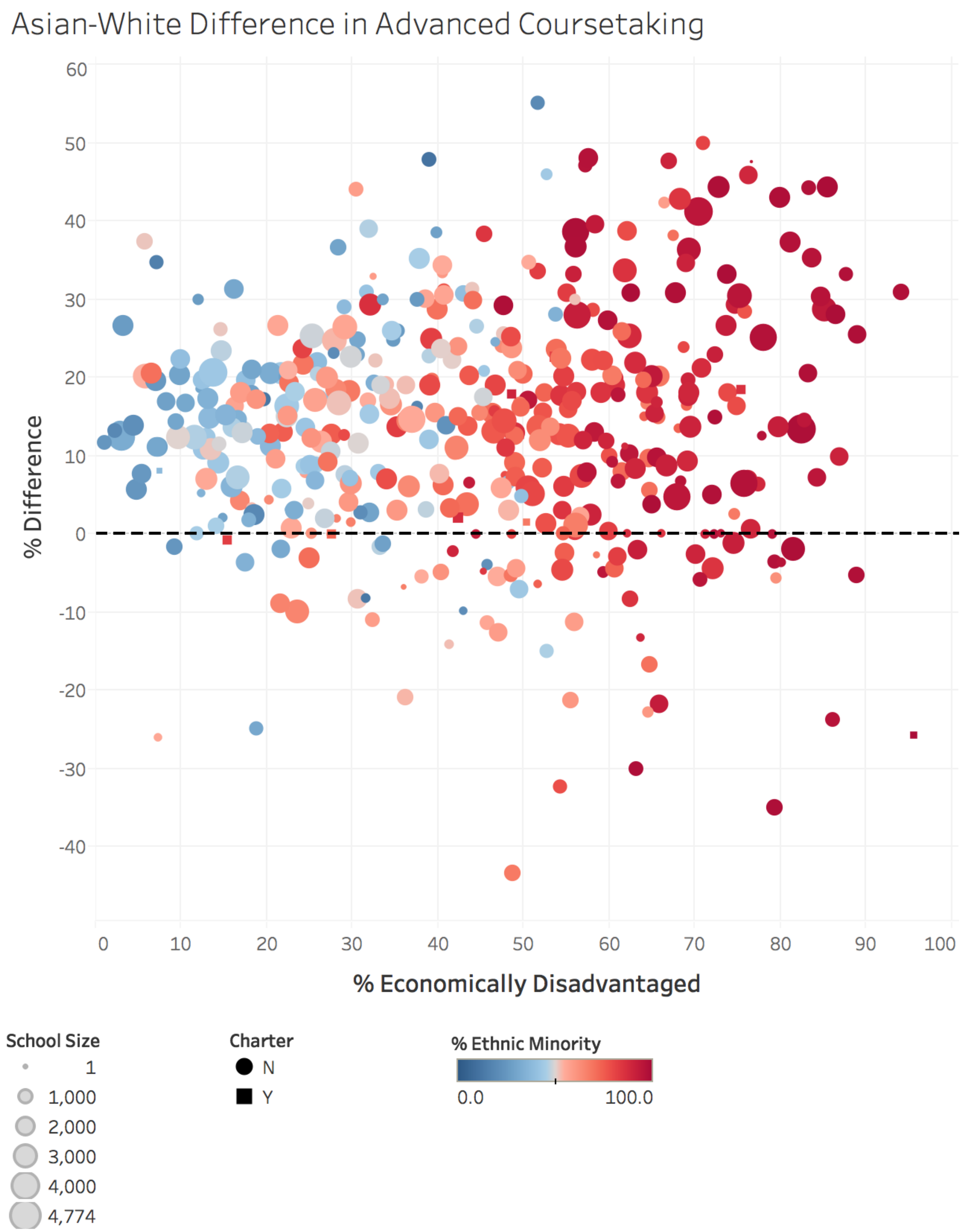


Figure 9. Difference in advanced course-taking between Black and White students in Texas public schools by percentage of economically disadvantaged students.

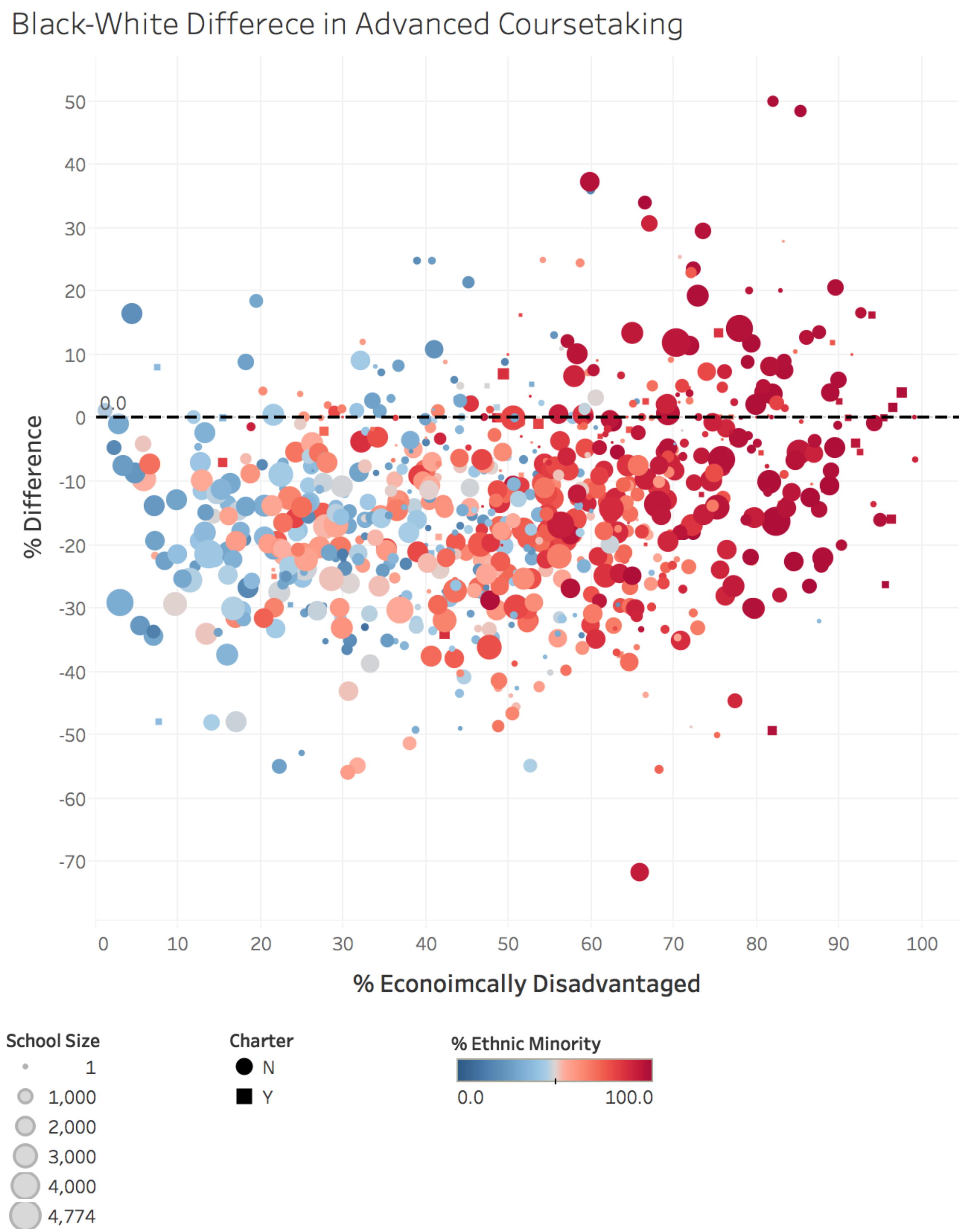
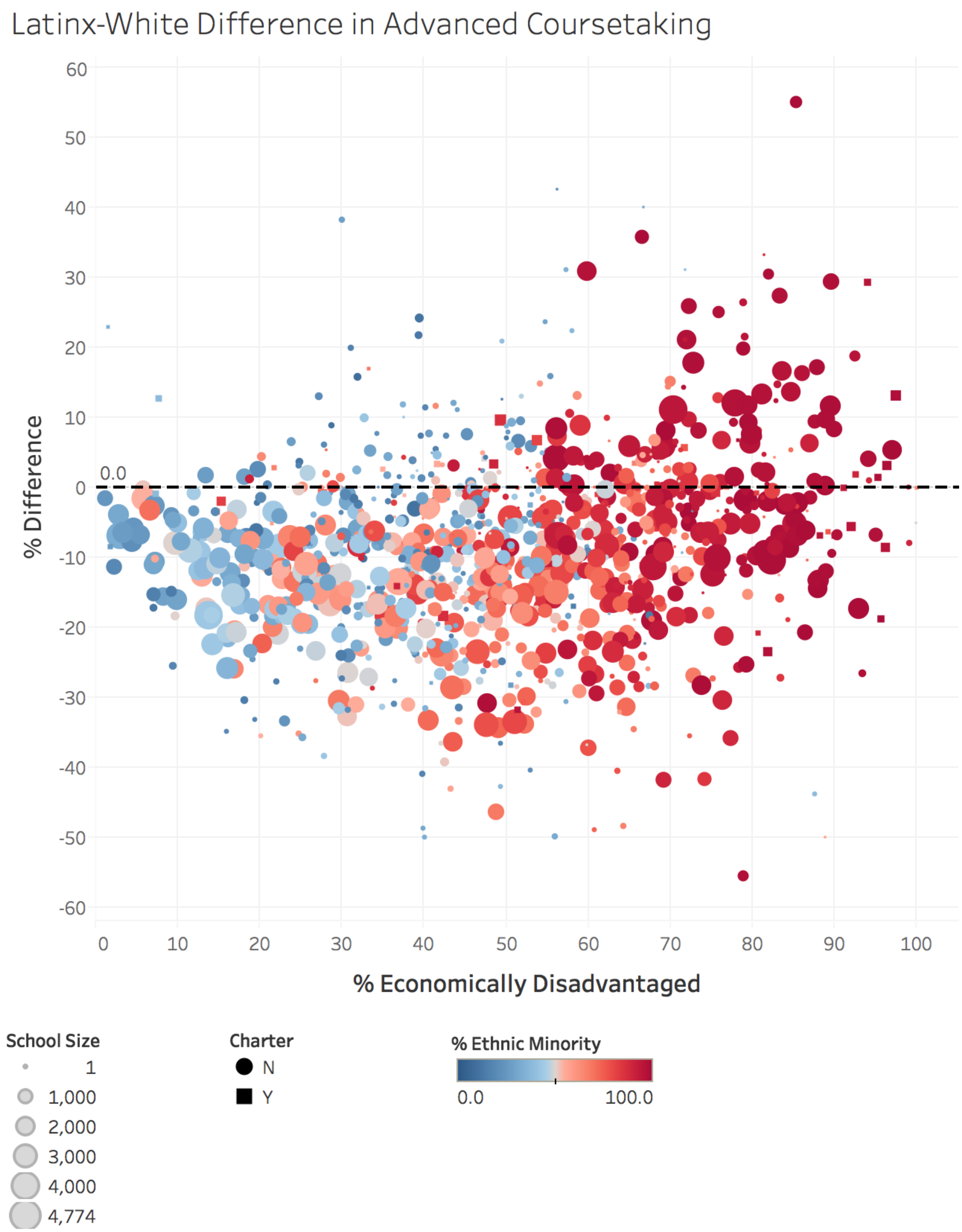


Figure 10. Difference in advanced course-taking between Latinx and White students in Texas public schools by percentage of economically disadvantaged students.



To interpret the plot, it is helpful to pay attention to the bolded line at the horizontal axis: data points falling above this line indicate that a higher percentage of Asian students enroll in advanced coursework than the percentage of all students enrolling in advanced coursework within a given school. As is evident in Figure 4, in which nearly all data points fall above the horizontal axis, Asian students are overrepresented in advanced coursework, a trend that seems to be independent of school level demographics or school level economic disadvantage.

Figure 5, Figure 6, and Figure 7 show similar plots for Black, Latinx, and White students, respectively. The interpretation of these plots follows the same logic as the interpretation of the plot in Figure 4. In Figure 5, nearly all schools fall below the horizontal axis, suggesting Black students are typically underrepresented in advanced coursework. Exceptions, however, should be noted in schools with high percentages of underrepresented minority students and with higher economic disadvantage. In several of these schools, Black students are more likely to be enrolled in advanced courses. Similarly, Figure 6 illustrates that Latinx students are typically underrepresented in advanced coursework, except in schools with higher percentages of underrepresented minority students and with higher economic disadvantage. By contrast, Figure 7 shows that White students are typically overrepresented in advanced course sequences.

The trend in Figure 7 is more variable in schools with high percentages of economically disadvantaged students. Such variability should be expected, as the percentage economically disadvantaged students is inversely proportional to the percentage of White students in a given school. With fewer White students, the difference in the percentage of White students and all students enrolling in advanced coursework within a school is sensitive to a single White student either enrolling or not enrolling in advanced coursework. Moreover, in low poverty schools in Texas, the percentage of students who are White approaches 100, which results in the data in Figure 7 converging to the horizontal axis as the percentage of economically disadvantaged students approaches zero. Similarly, high poverty schools in Texas have higher percentages of Latinx populations, resulting in the data in Figure 6 converging to the horizontal axis as the percentage of economically disadvantaged students approaches 100. The plots in Figure 8, Figure 9, and Figure 10 instead compare the percentages of ethnic minority students in Texas public schools (Asian, Black, and Latinx, respectively) to the percentage of White students enrolled in advanced courses. By using the percentage of White students enrolling in advanced coursework as a reference, it is easier to see that Asian students are more likely than White students to enroll in advanced coursework (Figure 8), a trend that does not vary by the concentration of economic disadvantage. Generally, fewer percentages of Black and Latinx students enroll in advanced coursework than White students (Figure 9 and Figure 10, respectively). In schools with higher concentrations economic disadvantage, however, the percentages of Black and Latinx students taking advanced coursework is greater than the percentage of White students taking advanced courses.

SOCIAL NETWORK ANALYSIS AND COMMUNITY DETECTION

In contrast to inferential statistics, which focuses upon attributes of individuals within a group (e.g., race and gender) to explore trends between these attributes and an outcome of interest, social network analysis uses the relationships between actors within a group to construct a network and explore how network characteristics explain the behavior of either the system or individuals and subgroups within the system (Borgatti & Ofem, 2010; Carolan, 2014). Networks can be constructed at a variety of grain-sizes (e.g., student, school, or district level analyses) or they can be constructed to explore relationships between actors at different levels (e.g., students and classes or friendships and extracurricular participation). The network perspective reflects a shift from traditional social science in which relationships between actors are used to contextualize actors' behaviors rather than quantifying the similarity between various actors based on their communal attributes. At an individual level, Katz, Lazer, Arrow, and Contractor (2004) explain, "people's behavior is best predicted by examining not their drives, attitudes, or demographic characteristics, but rather the web of relationships in which they are embedded. That web of relationships presents opportunities and imposes constraints on people's behavior" (p. 311-312).

A sociogram is a visualization technique used to show how actors within a network are related to one another. Sociograms consisting of one type of node (e.g., just students) are known as one-node networks, and sociograms consisting of two types of nodes (e.g., students and courses) are known as bipartite networks. Nodes (or vertices) within sociograms represent individual agents, and ties (or edges) between nodes indicate that a relationship exists between two actors. Edges can be either directed, indicating that one node interacts with another but not reciprocally (such as links from one webpage to another) or undirected, indicating that the edge represents a mutual interaction, such as two students who are in the same class. In addition, edges can be weighted according to the strength of the interaction between two nodes (Barrat, Barthélemy, Pastor-Satorras, & Vespignani, 2004; Lancichinetti & Fortunato, 2009).

Various network measures can be used to characterize the structure of the network, and lend insight into how the network structure influences individual actors' behaviors within the network (Borgatti & Ofem, 2010; Carolan, 2014; Katz et al., 2004). Examples of network characteristics include network density (Equations 5 and 6 in the "Community Detection in Networks" section), defined as the number of connections between nodes within a network divided by the total number of possible connections in that network, node degree (Equations 3 and 4 in the "Community Detection in Networks" section), the sum of the edges connected to a node, and betweenness centrality, a measure quantifying how much a given node connects other nodes in the network (Katz et al., 2004). In addition to network characteristics, it is possible to identify groups in social network analysis through community detection algorithms (Bruna & Li, 2017; Girvan & Newman, 2002; Newman, 2006; Newman & Girvan, 2004; Yang et al., 2016). Network characteristics or communities identified within the network can then be used as outcome or predictor variables in additional statistical analyses.

Constructing Sociograms

To construct a sociogram, an *m* by *n* matrix, **A**, is constructed in which matrix element $A_{vw} \in \{0, 1\}$ indicates whether or not node *v* is associated with event *w*. In this work, v is used to represent an individual student or school and w indicates either that student v was enrolled in course w or that school v offered course w. An example of such a matrix is given by Equation 1:

$$\boldsymbol{A} = \begin{bmatrix} A_{11} & \cdots & A_{1w} \\ \vdots & \ddots & \vdots \\ A_{v1} & \cdots & A_{vw} \end{bmatrix} \tag{1}$$

Matrix **A** produces a bipartite graph linking nodes from one level (i.e., students or schools) to nodes from a second level (i.e., STEM courses). Matrix **A** can also be multiplied by its transpose $\mathbf{A}^{\mathbf{T}}$, as in Equation 2, to produce a weighted adjacency matrix in which each element gives the edge weight, or the strength of the connection, between two nodes:

$$\begin{aligned} \boldsymbol{A} \cdot \boldsymbol{A}^{\boldsymbol{T}} &= \begin{bmatrix} A_{11} & \cdots & A_{1w} \\ \vdots & \ddots & \vdots \\ A_{v1} & \cdots & A_{vw} \end{bmatrix} \cdot \begin{bmatrix} A_{11} & \cdots & A_{v1} \\ \vdots & \ddots & \vdots \\ A_{1w} & \cdots & A_{vw} \end{bmatrix} \\ &= \begin{bmatrix} (A_{11} \cdot A_{11} + \cdots + A_{1w} \cdot A_{1w}) & \cdots & (A_{11} \cdot A_{v1} + \cdots + A_{1w} \cdot A_{vw}) \\ \vdots & \ddots & \vdots \\ (A_{v1} \cdot A_{11} + \cdots + A_{vw} \cdot A_{1w}) & \cdots & (A_{v1} \cdot A_{v1} + \cdots + A_{vw} \cdot A_{vw}) \end{bmatrix} \end{aligned} \tag{2}$$

The weighted adjacency matrix then produces a one-node network in which edges are weighted by the strength of the connection between two nodes (i.e., the number of STEM courses that either two students share or the number of STEM courses common between two schools). From the adjacency matrix, it is possible to uncover the underlying structure of the network by identifying smaller communities within that network. Methods for doing so are reviewed in the following section.

Community Detection in Networks

Communities within networks are defined as groups of nodes that are highly connected to other nodes within the community, but loosely connected to nodes outside of the community (Fortunato, 2010; Girvan & Newman, 2002; Reichardt & Bornholdt, 2006). Mathematically, this is represented by maximizing the difference between the internal and external densities of a community over all communities in a network (Fortunato, 2010). For a community $\mathcal{C}$ in network $\mathcal{N}$, the internal degree of a node v within $\mathcal{C}$, k_v^{int}, is defined as the sum of the edges between v and all other nodes in $\mathcal{C}$. The external degree of node v, k_v^{ext}, is the sum of the edges between v and all other nodes outside of $\mathcal{C}$. The total degree of node v is given by $k_v^{tot} = k_v^{int} + k_v^{ext}$ (Fortunato, 2010). The internal and external degrees of community $\mathcal{C}$ are equal to the sum of the internal and external degrees of the vertices within that community:

$$k_{\mathcal{C}}^{int} = \sum_{v} k_v^{int} \tag{3}$$

$$k_{\mathcal{C}}^{ext} = \sum_{v} k_v^{ext} \tag{4}$$

By the handshaking lemma (Wu, 2014), the number of internal and external edges of community $\mathcal{C}$ is then $\frac{k_{\mathcal{C}}^{int}}{2}$ and $\frac{k_{\mathcal{C}}^{ext}}{2}$, respectively, and the internal and external density of $\mathcal{C}$ is given by (Fortunato, 2010):

$$\delta_{int}(\mathcal{C}) = \frac{k_{\mathcal{C}}^{int}/2}{n_c(n_c-1)/2} = \frac{k_{\mathcal{C}}^{int}}{n_c(n_c-1)} \tag{5}$$

$$\delta_{ext}(\mathcal{C}) = \frac{k_{\mathcal{C}}^{ext}}{n(n-n_c)} \tag{6}$$

Maximizing $\sum_{\mathcal{C}}[\delta_{int}(\mathcal{C}) - \delta_{ext}(\mathcal{C})]$ over all communities within network $\mathcal{N}$ identifies community structure within a network.

A number of algorithms have been developed for uncovering the community structure within a network, many of which use physical analogies. The edge-betweenness algorithm, created by Girvan and Newman (2002), iteratively removes edges with high betweenness from a given network. Specifically, Girvan and Newman (2002) generalize betweenness centrality—a measure of how many paths intersect with a given node in a network—to edges, defining edge-betweenness as the number of pairs of nodes in a network that are connected through an edge. By iteratively removing edges with high betweenness, the edge-betweenness algorithm identifies groups of nodes that are highly interconnected, but loosely connected to nodes outside of these respective groups. Newman and Girvan (2004) describe different ways of calculating edge betweenness, including one method in which each edge within the network is considered to have a unit resistance, and betweenness is determined by summing over the voltage differences between all pairs of nodes.

Another community detection algorithm treats a network as a spin glass system in which couplings between nodes are ferromagnetic when they belong to the same community and antiferromagnetic when they belong to different communities (Reichardt & Bornholdt, 2006). The Hamiltonian of this system is given by Equation 7:

$$\mathcal{H} = -J\sum_{v,w} A_{vw}\delta(\sigma_v, \sigma_w) + \gamma\sum_{s}^{q}\frac{n_s(n_s - 1)}{2} \tag{7}$$

Here, A_{vw} is an element from the adjacency matrix indicating whether or not nodes v and w are connected; σ_v gives the spin state of node v; $\delta(\sigma_v, \sigma_w)$ is Kronecker's delta function which is 1 when $\sigma_v = \sigma_w$ and 0 when $\sigma_v \neq \sigma_w$; n_s is the number of spins in state s; and J and γ are coupling parameters that can be adjusted such that the entire network can be grouped into a single community or such that each node can consist of its own community. Fortunato (2010) notes that when the ratio γ/J is roughly equal to the average density of a network $\mathcal{N}$, i.e. when $\frac{\gamma}{J} \approx \delta(\mathcal{N})$, the energy of the spin glass system is minimized when spins align with communities where the internal and external densities are greater and less than, respectively, the average network density: $\delta_{int}(\mathcal{C}) > \delta(\mathcal{N})$ and $\delta_{ext}(\mathcal{C}) < \delta(\mathcal{N})$.

While many community detection algorithms draw inspiration from physical systems, several others try to maximize the modularity of a system. Modularity, given by Equation 8, compares the density of edges within and between communities in a network.

$$Q = \frac{1}{2m}\sum_{vw}\left[A_{vw} - \frac{k_v k_w}{2m}\right]\delta(c_v, c_w) \tag{8}$$

Here, A_{vw} is an element of the adjacency matrix and k_v is the degree of node v. Alternatively for weighted networks, A_{vw} gives the edge weight between nodes v and w, k_v is the sum of the weights of edges connected to node v, c_v is the community to which node v belongs, $\delta(c_v, c_w)$ equals 1 when $c_v = c_w$ and 0 when $c_v \neq c_w$, and $m = \frac{1}{2}\sum_{vw} A_{vw}$.

One algorithm, created by Newman (2004), starts with n communities, each corresponding to a single node, and combines smaller communities into larger ones. The algorithm only keeps merges that either give the biggest increase or smallest decrease in modularity. This algorithm runs over all merges, even when the change in modularity is zero, e.g. $\Delta Q = 0$. Clauset et al. (2004) improved upon Newman's approach with the fast-greedy algorithm, which is particularly quick at dealing with large networks. In the fast-greedy algorithm, modularity is maximized using a matrix with the change in modularity, ΔQ_{ij}, for all community pairs i and j; a max heap[3] that stores the maximum modularity change for each row of the matrix; and a vector with elements a_i. ΔQ_{ij} and a_i are defined in Equations 9 and 10:

$$\Delta Q_{ij} = \begin{cases} \frac{1}{2m} - \frac{k_i k_j}{(2m)^2}, for\ connected\ i\ and\ j \\ 0, for\ unconnected\ i\ and\ j \end{cases} \tag{9}$$

$$a_i = \frac{k_i}{2m} \tag{10}$$

The max heap is populated using ΔQ_{ij} and a_i and the largest value of ΔQ_{ij} is kept. This is repeated until there is only one community left.

The multi-level algorithm is another community detection method that maximizes network modularity (Blondel et al., 2008). In this algorithm, nodes are placed into arbitrary communities and then moved from these communities into adjacent communities. The resulting change in modularity is calculated using Equation 11:

$$\Delta Q = \left[\frac{\sum_i (w_i^{in} + 2k_i^{in})}{2m} - \left(\frac{\sum_i (w_i^t + k_i^t)}{2m}\right)^2\right] - \left[\frac{\sum_i w_i^{in}}{2m} - \left(\frac{\sum_i w_i^t}{2m}\right)^2 - \left(\frac{\sum_i k_i^t}{2m}\right)^2\right] \tag{11}$$

[3] A max-heap is a tree-like data structure in which parent nodes have a greater value than their child nodes.

In Equation 11, w_i gives the weights of edges connected to a node i, k_i gives the degree of node i, and these are summed over the total network (t) or community (in) to which a node belongs. If there is a positive change in modularity, the node is then moved into that community, and node weights are recalculated. These two steps are repeated until modularity is maximized.

Comparing Community Detection Algorithms

Testing the efficacy of various community detection algorithms requires that the results of these algorithms are compared to networks with known structure. Girvan and Newman (2002) introduced a benchmark in which computer generated networks have 128 nodes, 4 communities each with 32 nodes, and an average node degree of 16. Lancichinetti et al. (2008) and Lancichinetti and Fortunato (2009) note that this benchmark is relatively contrived, as real networks rarely exhibit the characteristics of the Girvan and Newman (GN) benchmark. Lancichinetti et al. (2008) propose a different benchmark based upon a mixing parameter, given by Equation 12:

$$\mu = \frac{\sum_i k_i^{ext}}{\sum_i k_i^{tot}} \tag{12}$$

Lancichinetti and Fortunato (2009) generalize the mixing parameter to weighted networks by defining node strength, s_i, as the sum of the edge weights adjacent to node i. The authors note that the strength of a node is related to its degree k_i through a power law:

$$s_i = k_i^{\beta} \tag{13}$$

The value of β in real networks is found to be $\beta \approx 1.5 \pm 0.1$ (Barrat et al., 2004). Lancichinetti and Fortunato (2009) define the weighted mixing parameter by Equation 14:

$$\mu_w = \frac{\sum_i s_i^{ext}}{\sum_i s_i^{tot}} \tag{14}$$

The mixing parameter takes a value between 0 and 1, corresponding to the ratio between the number of connections outside of the network to the number of connections within a network. When the mixing parameter takes values less than 0.5, a network is considered to have strong community structure.

Varying the mixing parameter and size of networks, Yang et al. (2016) compare the accuracy and computing time of various community detection algorithms available in the R package, igraph (Csardi, 2015). The accuracy of the community detection algorithms is measured by using the normalized mutual information (NMI) function, defined in Equation 15:

$$I(P,\bar{P}) = \frac{-\sum_i^C \sum_j^{\bar{C}} N_{ij}\, log\left(\frac{N_{ij}N}{N_{i0}N_{j0}}\right)}{\sum_i^C N_{i0}\, log\left(\frac{N_{i0}}{N}\right) + \sum_j^{\bar{C}} N_{0j}\, log\left(\frac{N_{0j}}{N}\right)} \tag{15}$$

In the NMI function, N_{ij} is an element of the confusion matrix **N**, in which rows correspond to real communities and columns correspond to the communities identified by a given algorithm. As such, N_{ij} gives the number of nodes belonging to "real" community i that also appear in community j uncovered by an algorithm. N gives the total number of nodes in the network, N_{i0} is a sum over the i^{th} row, N_{0j} is a sum over the j^{th} column, C corresponds to "real" communities, $\bar{C}$ corresponds to communities identified by an algorithm, P represents the partitioning of the real network, and $\bar{P}$ is the partitioning of the network as identified by the algorithm (Yang et al., 2016). A value of $I(P,\bar{P}) = 1$ indicates that an algorithm has perfectly identified the communities present in a real network.

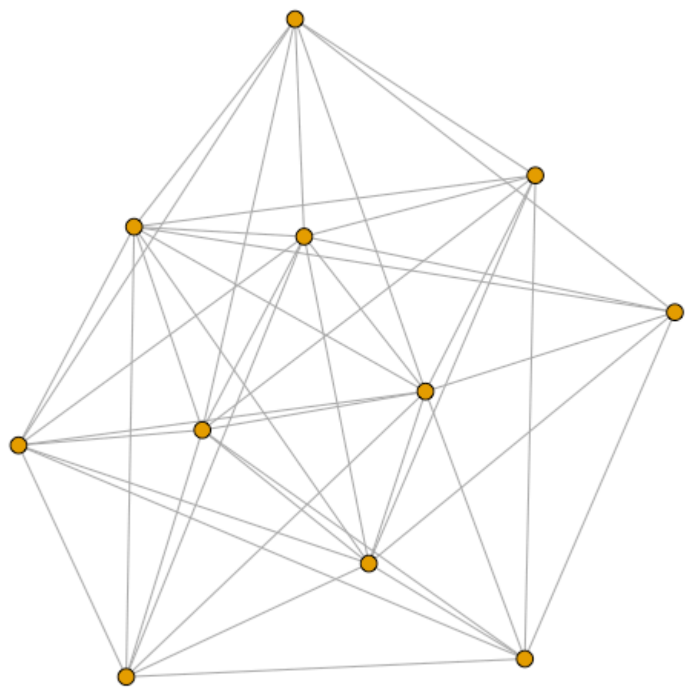
Using the NMI as a basis for the accuracy of community detection algorithms, Yang et al. (2016) recommend specific algorithms for networks given the network's size and the value of the mixing parameter of its community structure. Of note, the edge-betweenness and spin glass algorithms are computationally demanding and are not recommended for large networks. Of the algorithms that maximize modularity, the multilevel algorithm from Blondel et al. (2008) is found to be effective for networks with between 0 and 6,000 nodes and with a mixing parameter between 0.0 and 0.6. Given the versatility of the multilevel community detection algorithm and the characteristics of the networks explored in the present study, the multilevel algorithm is used in this analysis.

An important caveat to acknowledge is that for the sake of simplicity, the benchmarks against which community detection algorithms are measured are typically undirected, unweighted, and sparse, meaning the number of edges E is far less than the number of nodes N within a network. Fortunato (2010) notes that community detection is not possible on densely interconnected networks (i.e., when $E \sim N$) unless the edges connecting nodes within the network are heterogeneously weighted. The networks considered in this study are characterized by nodes with heterogeneously distributed weights, and community detection is therefore valid. A limitation, however, is that decisions regarding the specific algorithms employed are based off of studies that use unweighted, sparse networks as benchmark networks to test the efficacy of community detection algorithms.

Network Visualization

In addition to serving as the basis for many community detection algorithms, physical analogies are also often used to visualize networks. One such algorithm for determining network layouts is the Fruchterman-Reingold algorithm (Fruchterman & Reingold, 1990). The Fruchterman-Reingold algorithm is a force-directed graphing technique in which nodes are given an electrostatic charge such that all nodes experience repulsive forces between one another and edges between nodes act as springs that exert an attractive force between pairs of nodes (Fruchterman & Reingold, 1990). The nodes and forces between them form a system, which is then allowed to relax such to minimize the system's energy. An advantage of social network analysis is that it allows emergent trends within the data to be visually represented, leading to easier identification of these trends (Borgatti & Ofem, 2010; Carolan, 2014). The Fruchterman-Reingold algorithm leverages physical principles in order to generate a sociogram in which the proximity of nodes represents the strength of their relationship. An example of a sociogram using the Fruchterman-Reingold layout with simulated data is provided in Figure 11.

Figure 11. A sociogram using the Fructherman-Reingold algorithm to place nodes.



Chapter 3: Methods

In order to explore course offerings and student course-taking patterns in Texas charter and non-charter public schools, this study analyzes administrative educational data curated by the Texas Education Research Center (ERC). The Texas ERC collects and maintains: teacher certification data from the State Board for Educator Certification (SBEC); K-12 student-level demographic, enrollment, performance, and assessment data from the Texas Education Agency (TEA); campus and district level administrative data from the TEA; and post-secondary student-level demographic, enrollment, and performance data from the Texas Higher Education Coordinating Board (THECB). Texas began routinely collecting administrative educational data in the 1980's when the Texas Public Education Information Management System (PEIMS) was created. Files linking students to individual courses within schools were made available beginning in the 1993-1994 school year, and beginning in the 2011-2012 school year, files linking teachers to students and classes were also available.

The goal of this work is to explore differences in STEM course offerings and STEM course-taking patterns between charter and non-charter secondary schools and therefore uses the following data available through the Texas ERC: student demographic data (e.g., race, gender, and designation as economically disadvantaged, limited English proficiency (LEP), special education (SPED), and gifted); student secondary coursework in STEM; school sector (charter or non-charter public school); and school level demographic data (which are obtained by aggregating student level demographic data at the campus level). To create a data set for this analysis, students are matched to STEM courses for each school in Texas by year. A state-wide high-school cohort is constructed by identifying all Texas students beginning 9th grade for the first time in the 2011-2012 school year. The students in this cohort are then followed over four years, the standard time to complete high school, creating a data file with student identification, school identification, unique course identification by year (e.g., adv_physics_2012 to denote that a student took advanced physics during the 2011-2012 school year), and student demographic data.

Only schools that are open during all four years for a given cohort are included in the data set. Including schools that are only open for a fraction of a cohort's four years may serve to bias the results, as these schools do not offer four full years of STEM coursework to students. Moreover, schools classified as Disciplinary Alternative Education Programs (DAEP) or Juvenile Justice Alternative Education Programs (JJAEP), which serve students who have been removed from the schools due to felonious activity, are not included in the present study. The course offerings in these schools are limited and the number of students completing four years in DAEP and JJAEP schools is relatively few. Moreover, the academic programs offered in these schools do not necessarily reflect sector differences but are instead established to meet the needs of students who have faced severe disciplinary action within the public education system. These schools are not included because they are characteristically not emblematic of school choice. Given these decisions, a total of 1630 charter and non-charter secondary schools are included in the present study. The average demographic characteristics of charter and non-charter schools included in this study are provided in Table 2:

Table 2. Average demographic characteristics of Texas charter and non-charter secondary schools.

	Charter	Non-charter
# of Schools	178	1452
Cohort Size	51	208
% FRL	67.0	52.6
% LEP	7.9	4.1
% SPED	11.4	10.8
% Gifted	3.1	9.2
% Female	52.1	48.4
% Asian	2.1	1.8
% Black	15.5	10.3
% Latinx	58.2	42.0
% White	22.3	43.6
% 8^{th} Advanced Math	19.7	19.2

In addition to removing schools that are not open for all four cohort years and DAEP/JJAEP schools, students who leave a school for reasons other than dropping out or transferring to another school are removed. Students who exit a school, but neither drop out of high school entirely nor transfer to another school, exit for reasons that do not necessarily reflect their academic trajectory (e.g., returning to a students' home country, death, homeschooling, or moving out-of-state). Moreover, these exit reasons do not necessarily preclude students from continuing or discontinuing a certain curricular path (i.e., students' out-of-state course-taking is unavailable in Texas ERC data).

To track students who either drop out of the public-school system entirely or transfer to another school, dummy "course" variables are created in order to indicate the year and manner in which a student leaves a school (e.g., transfer_out_2012 indicates that a student transferred from a school in 2012). Student exit (either drop-out or transfer) in addition to student STEM course enrollment comprise a student's secondary course memberships from which a network is created.

In order to account for students enrolled in more than one campus in a given academic year or who transfer to another campus in a subsequent academic year, a students' membership in the data set is weighted by 0.25 for each year in which that student is present. This weight is then divided by the number of schools attended by a given student in that year. Because students are followed across the four years after which they begin high school, these weights are used to compute cohort-level demographics in schools. A school cohort is defined as the total group of students who attended that school for either the entirety of their high school careers or a fraction thereof. In addition to using these weights to compute cohort demographics, student weights are used in student-level statistical models, described in greater detail later in this chapter.

This study analyzes course-offerings and course-taking in Texas charter and non-charter school at three different levels. A state-wide analysis is first conducted to provide an overview of STEM course-taking in Texas and to contextualize subsequent analyses. Then school-level analyses are conducted in which STEM course-offerings in charter and non-charter schools are compared. Finally, student-level analyses are conducted in order to investigate prominent STEM course-taking patterns among students in Texas charter and non-charter secondary schools. The methods used in each of these analyses are described in the following subsections.

Texas Statewide STEM Course-Taking

To examine statewide STEM course-taking patterns in Texas, an adjacency matrix of STEM courses offered in Texas across the four years of this study is created in which the elements of the matrix give the number of students shared between two courses. This is done by creating a student-course matrix, $\mathbf{A}$, in which matrix element A_{vw} indicate that student v was enrolled in course w. The matrix product, $\mathbf{A^T} \cdot \mathbf{A}$, gives the adjacency matrix in which edges are weighted by the number of students common to two STEM courses in Texas. The resulting sociograms are plotted using the Fruchterman-Reingold algorithm (1990). To identify which courses are most closely associated with one another, the multilevel community detection algorithm is applied to the statewide course network (Blondel et al., 2008). Course identifiers used in the present study also include the year in which a given STEM course was offered, so the statewide analysis provides an overview of the sequence and content of courses that are associated with one another in Texas. In addition to looking at the entirety of course-taking across Texas, course-taking is also analyzed by sector, using the Fruchterman-Reingold algorithm (1990) to display course networks for charter and non-charter schools in Texas. The goal of the statewide analysis is primarily descriptive in order to provide context for the remaining analyses conducted in this study.

STEM Course-Offerings in Charter and Non-Charter Schools

At the school level, matrix $\mathbf{A}$ in which element A_{vw} indicates whether or not school v offers course w is multiplied by its transpose, $\mathbf{A} \cdot \mathbf{A^T}$, in order to create a weighted adjacency matrix in which matrix elements are weighted by the number of courses common to each pair of schools. As with the statewide STEM course-taking analysis, school-level networks are visualized using the Fruchterman-Reingold algorithm (1990), and the multilevel community detection algorithm is applied to identify communities of schools in the network due to common course offerings (Blondel et al., 2008). The course offerings in each community of schools are then analyzed to determine the scope of STEM course-offerings in that community.

After identifying the courses most commonly shared between schools within a community, a multinomial logistic regression model (Equation 16) is used to determine the odds of a school belonging to community α relative to a reference community given school level predictor variables:

$$log\left(\frac{\theta_i^\alpha}{\theta_{ref}}\right) = \beta_C C_i + \beta_S S_i + \beta_M M_i + \sum_i \beta_X X_i \quad (16)$$

Here, $\frac{\theta_i^\alpha}{\theta_{ref}}$ gives the odds ratio of campus i belonging to community α relative to a reference community, C_i indicates whether or not a school is a charter school, S_i gives a school's size as determined by its cohort population, M_i gives the proportion of students at school i who enrolled in a high-school level math course (at least algebra 1) in 8th grade, and X_i are school level demographic variables for the percentages of economically disadvantaged, special education (SPED), underrepresented minority, gifted, and limited English proficient (LEP) students.

School level analyses are also conducted using results from student-level community detection (explained in the following section). In the student-level community detection, communities of students as a function of the number of courses common to these students are identified for each school. The student-level communities identified within Texas charter and non-charter schools are then categorized using a k-means clustering algorithm. Two school-level analyses are conducted using these results. One analysis, given by Equation 17, seeks to determine whether the number of communities (e.g., course-taking pathways) differs between charter and non-charter schools:

$$P_i = \beta_C C_i + \beta_S S_i + \beta_M M_i + Dist_{j[i]} + \sum_i \beta_X X_i \quad (17)$$

In Equation 17, P_i gives the number of pathways identified within school i, and the predictor variables are the same as those used in Equation 16, except that a random effects variable for district j to which school i belongs is also included. A regression analysis identical to Equation 17, but without the district random effect coefficient, is also run, but no difference in the statistical significance between the predictor variables is found.

Finally, a multinomial logistic regression model similar to the model specified by Equation 16 is run in which the course-taking pathways identified during student-level community detection analysis serve as the outcome communities. In addition, this model also controls for the number of pathways P_i identified in school i:

$$log\left(\frac{\theta_i^{\alpha}}{\theta_{ref}}\right) = \beta_C C_i + \beta_S S_i + \beta_M M_i + \beta_P P_i + \sum_i \beta_X X_i \tag{18}$$

STUDENT STEM COURSE-TAKING IN CHARTER AND NON-CHARTER SCHOOLS

The final level of analysis included this study is a student-level analysis. As with the school-level and state-level analysis, a student-course matrix **A** is created in which element A_{vw} indicates whether or not student v is enrolled in course w for each school included in the study. A weighted adjacency matrix is constructed by multiplying **A** by its transpose, and the elements of the resulting matrix are weighted by the number of courses within a school that each pair of students have in common. Due to the fact that these networks are at the student level, producing and displaying sociograms violates FERPA regulations. For each school, the multilevel community detection algorithm is used to find groups of students with common STEM course pathways throughout their high school careers (Blondel et al., 2008).

The average number of communities identified in Texas charter and non-charter secondary schools is about 3.4. Given that this study includes a total of 1630 schools, categorizing the communities of students based upon their shared courses requires analyzing course-taking in over 4,800 communities. k-means clustering is a method of partitioning data into groups in which the data points in each group have similar means across a vector of attributes (Guthrie, 2018; MacQueen, 1967). The k-means clustering algorithm takes a set of *N* data points, each with a vector of attributes $\overrightarrow{a_n}$, and partitions these data points into *k* groups: $S = (S_1, S_2, \ldots, S_k)$. Each group S_i is constructed such that the within-group sum of squares is minimized, as given by Equation 19:

$$\arg\min_{S} \sum_{i}^{k} \sum_{\vec{a} \in S_i} |\vec{a} - \overrightarrow{\mu_i}|^2 \tag{19}$$

In the present study, k is set to 6, thus identifying six distinct course-taking patterns in Texas charter and non-charter schools. The probability of a student belonging to a community characterized by one of the six course-taking patterns identified through k-means clustering is assessed through the multilevel logistic regression model specified by Equation 20:

$$log\left(\frac{\theta_i^{\alpha}}{\theta_{ref}}\right) = \beta_c C_i + \beta_M M_i + School_{j[i]} + \sum_{X} \beta_X X_i \tag{20}$$

In Equation 20, θ_i^{α} represents the odds that student *i* enrolls in course-sequence α, θ_{ref} represents the odds of enrolling in a reference course-sequence, C_i denotes whether or not student *i* attends a charter school, M_i is a flag indicating whether or not a student *i* took an advanced math course in middle school, and X_i are student level demographic variables for ethnicity and designation as SPED, LEP, economically disadvantaged, and gifted. In addition to these fixed effects, random effects coefficients for school *j* in which student *i* is enrolled are included in order to account for the fact that students in a school are not independent of one another. The outcome variable is the log odds ratio giving the likelihood of a student enrolling in a given course-sequence within a school as compared to a reference course-sequence.

Mixed-effects multinomial logistic regression models cannot be run in R; however, individual mixed-effects logistic regression models can be run using the package lme4. Begg and Gray (1984) show that it is possible to estimate a mixed effects multinomial regression model by running individual mixed-effects logistic regression models comparing each category of the outcome variable. Of note, this approximation is more conservative than a single multinomial regression model and works best when the reference category is the most common category of the outcome variable. Given that this work identifies 6 prominent course-taking patterns in Texas charter and non-charter schools, fifteen logistic regression models are run in order to compare the likelihood of a student taking one of the six identified course-sequences relative to the other five:

$$C_2^6 = \binom{6}{2} = \frac{6!}{2!\,4!} = 15 \tag{21}$$

Chapter 4: Results

The results of this study are presented according to the various levels of analyses conducted to investigate STEM course-offerings and course-taking patterns in Texas charter and non-charter secondary schools. First, statewide results are described, followed by results from school-level and student-level analyses.

TEXAS STATEWIDE STEM COURSE-TAKING

Sociograms of STEM courses in which edges are weighted by the number of all Texas schools offering each pair of courses are given in Figure 12 and Figure 13. Figure 12 is colored by the course category, in which grey indicates that the course is associated with an exit (transfer into, transfer out of, or dropping out of a school), dark blue nodes

Figure 12. Sociogram of STEM courses connected by all Texas public schools and colored by course category.

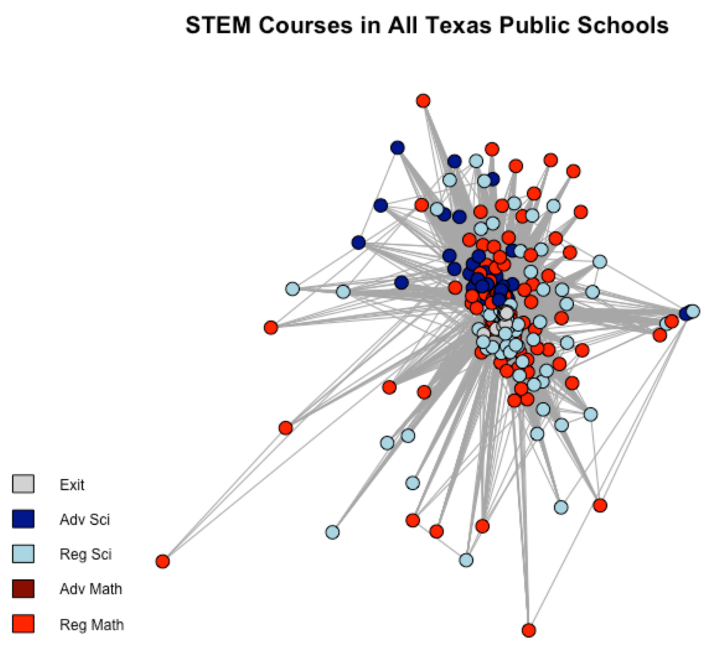
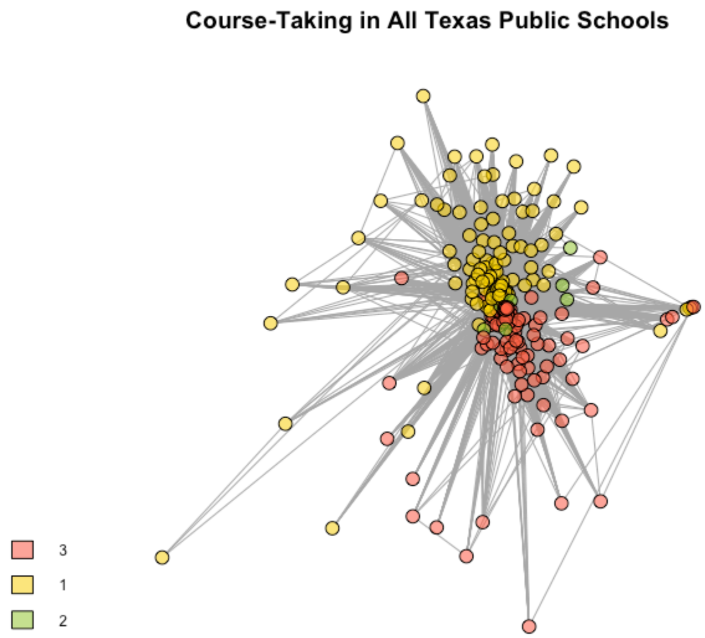


Figure 13. Sociogram of STEM courses connected by all Texas public schools and colored by multilevel community.



represent advanced science courses, light blue nodes represent regular science courses, dark red nodes represent advanced mathematics courses, and light red nodes represent regular mathematics courses. The multilevel community algorithm detected three communities in the network consisting of courses linked by all Texas public schools. Figure 13 is colored by these communities.

Sociograms displaying STEM courses with edges weighted by the number of Texas charter and non-charter public schools offering each pair of courses are similarly created. These sociograms are colored by course category (Figure 14 and Figure 16 for charter and non-charter schools, respectively) and community (Figure 15 and Figure 17 for charter and non-charter schools, respectively) detected using the multilevel algorithm. Four communities were detected using charter schools and three communities using non-charter schools.

Figure 14. Sociogram of STEM courses connected by Texas charter schools and colored by course category.

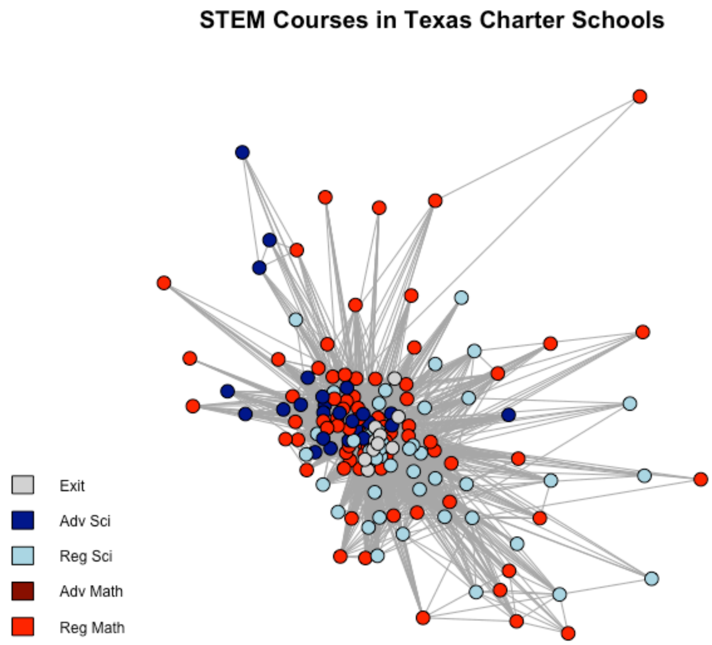


Figure 15. Sociogram of STEM courses connected by Texas charter schools and colored by multilevel community.

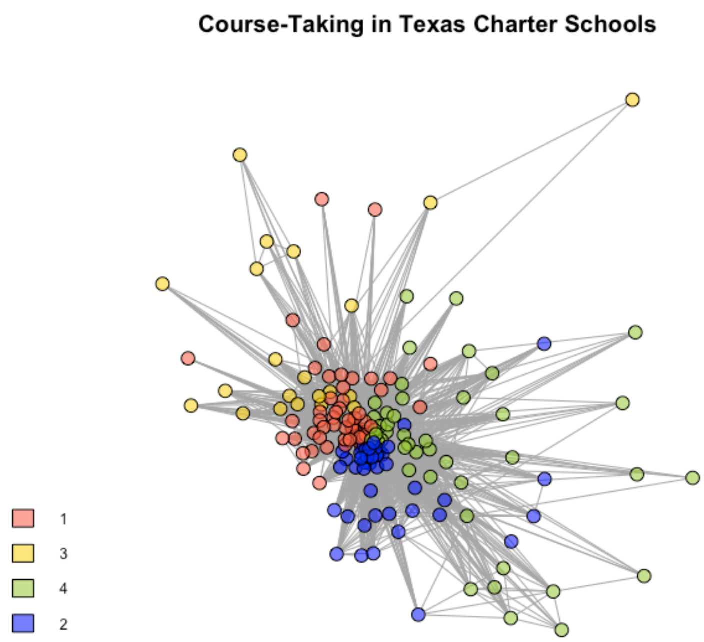


Figure 16. Sociogram of STEM courses connected by Texas non-charter schools and colored by course category.

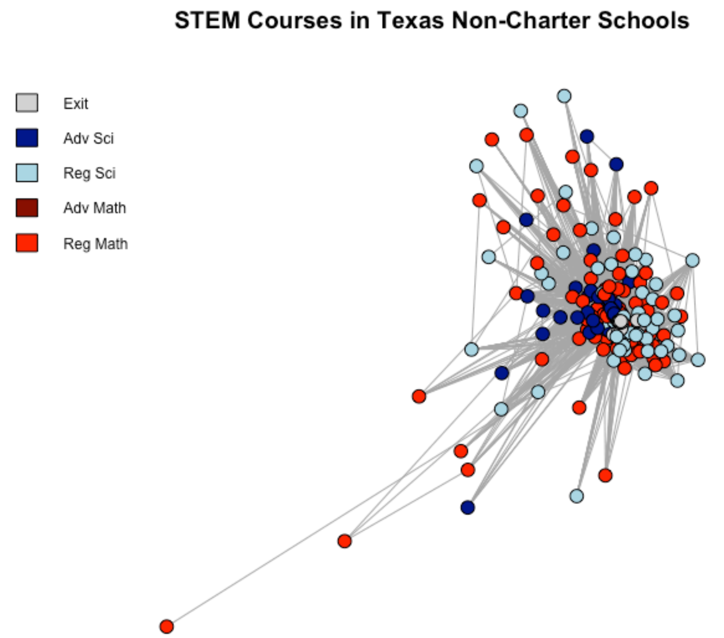
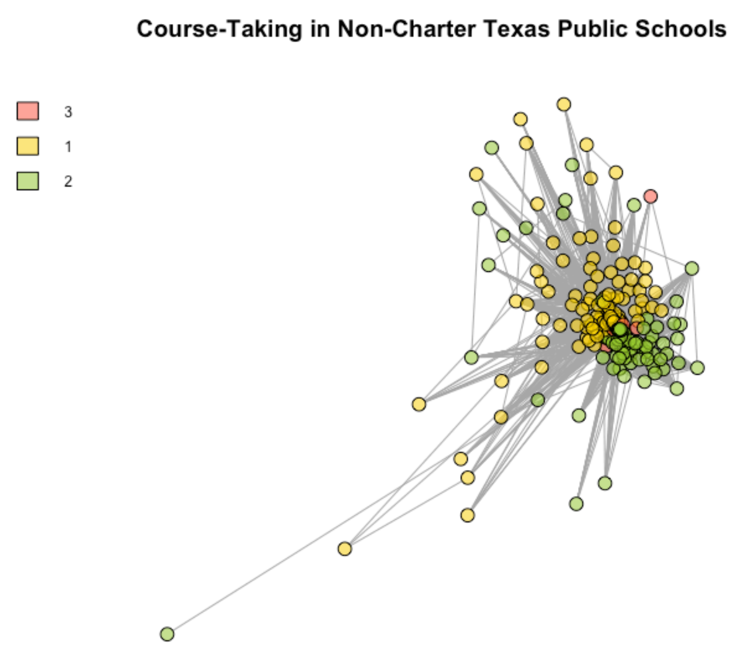


Figure 17. Sociogram of STEM courses connected by Texas non-charter schools and colored by multilevel community.



The following trends emerge after analyzing the STEM courses and years associated with the communities detected through the multilevel community detection algorithm. Community 1 detected in the network of all Texas public schools is characterized by rudimentary courses taken in later years (e.g., algebra 1 taken in 2014, when most students are in 11th grade) in addition to special education courses in algebra 1, algebra 2, geometry, biology chemistry, and physics taken over the entirety of the four-year time span during which this cohort is analyzed. Community 2, by contrast, is characterized by a number of AP courses taken at various points throughout the four-year time span. Community 3 consists of a "staple" STEM course-taking pattern (e.g., algebra 1, algebra 2, geometry, precalculus, biology, physics, and chemistry) supplemented with non-AP elective courses.

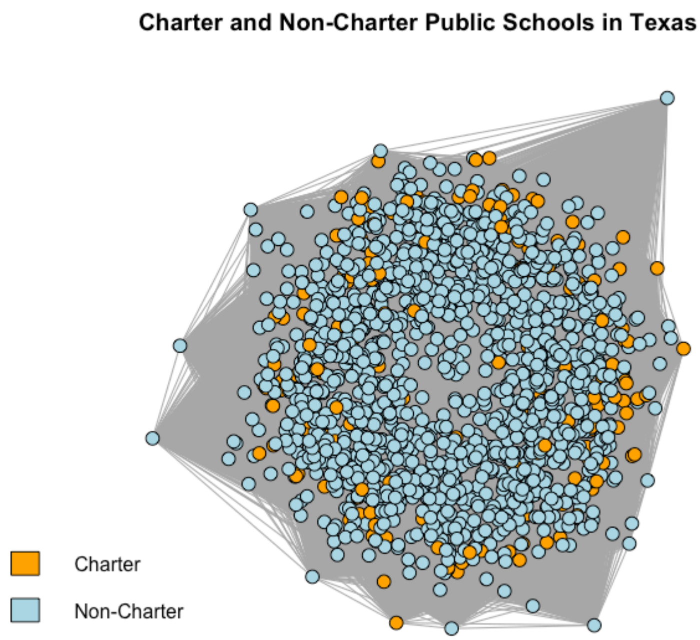
Community 1 detected in the network using Texas charter schools is characterized by STEM staples (e.g., algebra 1, algebra 2, biology, chemistry, etc.) taken early in the four-year high school period with AP and IB courses taken in later years. Community 2 is slightly more advanced than community 1, as AP and IB courses are taken earlier in the four-year high school period. Community 3 is similar to community 3 in the all schools network in that it includes the "staple" STEM curriculum with non-AP electives. Finally, community 4 for charter schools is similar to community 1 of the all Texas schools network in that it consists of several SPED courses and rudimentary courses in STEM taken later in the high school career.

In the non-charter school network, community 1 is characterized by the staple STEM curriculum with non-AP electives, community 2 is characterized by the staple STEM curriculum with AP courses taken throughout the high school curriculum, and community 3 consists of SPED courses and rudimentary courses taken later in students' high school careers.

STEM Course-Offerings in Charter and Non-Charter Schools

Sociograms in which nodes represent Texas public schools and edges are weighted by the number of STEM courses shared between pairs of schools are provided in Figure 18

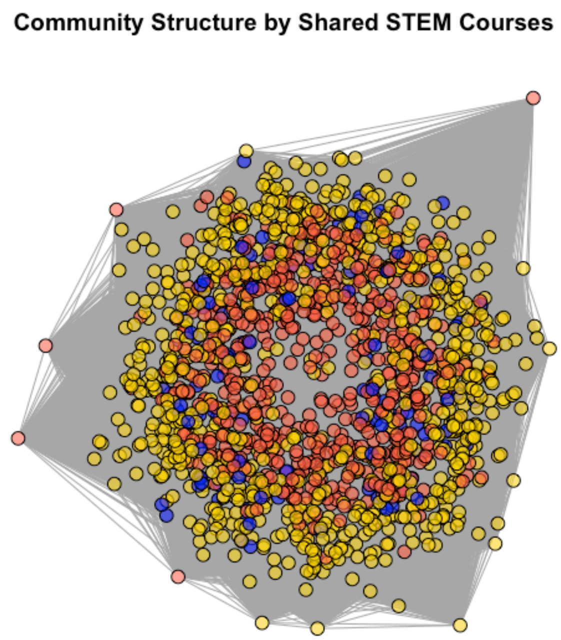
Figure 18. School-level sociogram in which edges are weighted by the number of courses shared by pairs of schools. The sociogram is colored by school sector.



and Figure 19. Figure 18 is colored by school sector, and Figure 19 is colored by the communities detected using the multilevel community detection algorithm. Community 1 (C1) consists of 641 schools, community 2 (C2) consists of 879 schools, and community 3 (C3) consists of 110 schools. The courses associated with the communities identified from the school-level network are listed in Table 3. Courses are grouped according to the percentage of schools in a given community that share those specific courses.

Schools in C1 have the most expansive course offerings, as they offer "staple" STEM courses in addition to those tailored for SPED students. Moreover, these schools offer a number of advanced courses. Schools in C2 seem to have the most restrictive STEM course offerings, characterized by the "staple" STEM curriculum with few advanced course offerings and few electives. In addition, schools in C2 do not offer STEM courses

Figure 19. School-level sociogram in which edges are weighted by the number of courses shared by pairs of schools. The sociogram is colored by community.



tailored for SPED students with a high frequency. Finally, C3, which represents the smallest community, offers “staple” STEM courses both for SPED students and students not designated as SPED. As with C2, C3 does not offer many advanced courses.

Coefficients from the multinomial logistic regression model specified in Equation 16 are given in Table 4. For this model, the probability of a school belonging to C1 serves as the reference category to which the probability of a school belonging either to C2 or to C3 is compared. Exponentiating the coefficients in Table 4 indicates how much a predictor variable increases or decreases the odds of a school belonging to a community (in our case, C2 or C3) relative to the reference community (C1).

Table 3. Courses associated with communities identified through school-level community detection.

% Schools	Community 1 (C1) 641 Schools	Community (C2) 879 Schools	Community 3 (C3) 110 Schools
75-100	Algebra 1 (SPED) Algebra 1 Algebra 2 AP Biology AP Calculus AB AP Chemistry Biology (SPED) Biology Chemistry Geometry (SPED) Geometry Integrated Phys/Chem Mathematical Models Physics Precalculus Dropout Transfer In Transfer Out	Algebra 1 Algebra 2 Biology Chemistry Geometry Integrated Phys/Chem Mathematica Models Physics Precalculus Transfer In Transfer Out	Algebra 1 (SPED) Algebra 1 Algebra 2 (SPED) Biology (SPED) Biology Chemistry Geometry (SPED) Geometry Integrated Phys/Chem Mathematical Models Physics Precalculus Transfer In Transfer Out
50-75	Advanced Quant Reasoning AP Calculus BC AP Environmental Science AP Physics 1 AP Statistics Environmental Science Independent Study, Math	Algebra 1 (SPED) Dropout Geometry (SPED)	Dropout Environmental Systems Independent Study, Math
25-50	Algebra 2 (SPED) AP Physics B AP Physics C Aquatic Science Astronomy Chemistry (SPED) Earth & Space Science	AP Calculus AB Biology (SPED) Environmental Systems Independent Study, Math	AP Calculus AB Earth & Space Science

With the exception of school size, which is in units of individual students, the other predictor variables in Table 4 represent the proportion of students classified as belonging to a given demographic predictor category at each school. Controlling for school level demographics and school size, a charter school is no more or less likely to belong to C2 relative to C1. Charter schools are, however, 93% less likely to belong to C3 relative to C1. In addition, schools with greater student populations and greater percentages of students who enrolled in a high-school level math course in 8th grade are less likely to belong to either C2 or C3 relative to C1. By contrast, increases in the percentages of Black, Latinx, multi-racial, and white students are associated with increases in the likelihood that a school belongs to either C2 or C3 relative to C1.

Table 4. Coefficients for the multinomial regression model specified in Equation 16.

	Log-odds, C2/C1			**Log-odds, C3/C1**		
	Est.	*SE*	*Sig.*	*Est.*	*SE*	*Sig.*
Charter	0.44	0.36		-2.70	1.11	*
Econ. Dis.	1.32	0.79	.	1.13	0.96	
SPED	-5.12	0.86	***	-1.39	0.90	
Gifted	-0.77	1.11		-1.52	1.81	
LEP	0.79	1.93		-3.98	3.17	
MS Math	-2.27	0.60	***	-3.40	0.99	***
Asian	1.01	2.07		9.88	2.20	***
Black	4.46	0.92	***	4.45	1.05	***
Latinx	5.28	0.78	***	5.15	0.81	***
Mutli-racial	9.49	2.03	***	13.13	1.73	***
Nat. Am.	0.53	2.80		3.83	2.40	
Pac. Is.	-28.85	0.05	***	-48.12	0.02	***
White	6.66	0.77	***	7.02	0.76	***
School Size	-0.03	0.002	***	-0.01	0.001	***
*** p < 0.001, ** p < 0.01, * p < 0.05, . p < 0.1						

Coefficients from the regression model predicting the number of course-taking pathways on account of school level predictors and charter school status (Equation 17) are provided in Table 5. The scales of the predictor variables in this model are the same as the scale of the predictor variables in the multinomial model just discussed. In the model specified by Equation 17, there is no statistically significant effect of charter school status on the number of course pathways identified in Texas public schools. Increasing school size and the percentages of Asian, Black, Latinx, Multi-racial, Native American, and White students in a school are associated with statistically significant increases in the number of course pathways identified in that school. Interestingly, increases in the percentages of gifted students and students who enrolled in a high school math course during 8th grade are associated with decreases in the number of course pathways identified in a given school.

Table 5. Regression coefficients for the model specified in Equation 17.

	Est.	*SE*	*Sig.*
Charter	-0.10	0.07	
Econ. Dis.	0.10	0.15	
SPED	-0.04	0.21	
Gifted	-1.14	0.29	***
LEP	0.69	0.35	
MS Math	-0.42	0.15	**
Asian	2.89	0.50	***
Black	3.92	0.20	***
Latinx	3.30	0.15	***
Mutli-racial	3.09	0.98	**
Nat. Am.	7.00	1.69	***
Pac. Is.	7.11	4.38	
White	3.07	0.08	***
School Size	0.001	0.0001	***
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$			

The third model used to evaluate course-offerings in Texas charter and non-charter schools uses communities identified through the student-level network, in which edges between nodes represent the number of courses common between pairs of students within each school in this study, as the outcome variable. As such, results from this model will be discussed in the following section.

STUDENT STEM COURSE-TAKING IN CHARTER AND NON-CHARTER SCHOOLS

As described earlier, the average number of communities identified in Texas charter and non-charter public schools was approximately 3.4. With 1630 schools in the present analysis, an average of 3.4 communities per school results in over 4800 communities identified at the student level. In order to group these communities together, k-means clustering was used to find clusters of communities in which the mean sum of squares for the vectors of attributes for each community was minimized (Equation 19). The following parameters were used to identify clusters of communities:

- the average number of total STEM courses taken by students;
- the average number of AP and IB courses taken by students;
- the average number of SPED courses taken by students;
- the average number of "staple" STEM courses taken by students[4];
- the average number of advanced courses taken by students, which is not limited to AP and IB coursework;
- the average total number of math and science courses taken by students;
- the average percentage of students with enough STEM credits to graduate under the distinguished academic program (DAP);
- the average percentage of students with enough STEM credits to graduate under the foundations graduation plan;
- the average percentage of students who dropped out;
- the average percentage of students who transferred out of a school;
- the average percentage of students who transferred into a school;
- the average percentage of students taking at least one AP or IB STEM course;
- the average percentage of students taking at least on advanced STEM course;
- and average percentage of students taking at least on SPED STEM course.

Using k-means clustering, 6 distinct types of course-pathways were discovered: an advanced pathway (Adv); a basic pathway (Basic); a college preparatory pathway (Cprep); a transition pathway (Trans); a SPED pathway (SPED); and an exit pathway (Exit). The

[4] The "staple" stem courses are defined as algebra 1, algebra 2, geometry, precalculus, biology, chemistry, and physics. These courses are often offered as *staples* of a college preparatory curriculum.

average values of the attributes used to identify these clusters of communities using the k-means algorithm are given in Table 6.

Table 6. Average values of parameters used to cluster communities using the k-means algorithm.

	Adv	Basic	CPrep	Trans	SPED	Exit
Avg. # STEM Courses	7.6	5.4	6.8	3.8	4.44	2.0
Avg. # Science Courses	3.8	2.5	3.5	1.4	2.4	0.7
Avg. # Math Courses	3.7	2.3	3.1	1.4	1.8	0.5
Avg. # AP/IB Courses	1.4	0.1	0.2	<0.1	<0.1	<0.1
Avg. # Adv. Courses	2.3	0.3	0.9	0.1	<0.1	<0.1
Avg. # "Staple" Courses	5.6	3.9	5.7	2.3	0.3	0.5
Avg. # SPED Courses	0.2	0.2	0.1	0.2	3.6	0.4
Avg. % Stu, DAP	48.8	9.6	29.5	<0.1	1.4	<0.1
Avg. % Stu, Foundations	15.5	13.2	15.7	<0.1	5.8	<0.1
Avg. % Stu, Drop	0.4	4.3	1.5	7.6	3.7	10.6
Avg. % Stu, Transfer In	9.1	30.9	14.5	42.9	10.2	52.3
Avg. % Stu, Transfer Out	8.1	28.9	10.1	51.4	11.2	14.2
Avg. % Stu, AP/IB	68.6	4.4	13.9	0.6	0.1	1.4
Avg. % Stu, Adv.	90.3	22.2	62.6	6.0	2.6	6.2
Avg. % Stu, SPED	1.3	8.5	2.6	8.0	96.6	22.2

Clusters of communities identified through the k-means clustering algorithm are used as the outcome variable in a school-level multinomial logistic regression model (Equation 18) and a series of student-level hierarchical logistic regression models (Equation 20). The cluster of communities with the largest number of students is the "College Preparatory" community and is therefore set as the reference category in both models. Average demographic characteristics of students in each of identified clusters of communities for charter, non-charter, and all Texas public schools are provided in Table 7. The coefficients from the school-level multinomial logistic regression model are provided in Table 8, and the coefficients from the fifteen student level hierarchical logistic regression models are provided in Table 9.

Table 7. Average demographic characteristics of students associated with clusters of communities identified through k-means clustering by school sector.

		Adv	Basic	CPrep	Trans	SPED	Exit
Texas Non-Charter Schools	Avg. % Econ. Dis.	31.7	60.1	54.8	70.0	73.1	64.5
	Avg. % LEP	1.3	7.7	5.7	10.3	7.9	4.7
	Avg. % SPED	0.9	15.7	5.2	18.9	91.3	21.1
	Avg. % Gifted	28.5	3.9	7.2	3.2	0.5	1.7
	Avg. % Female	51.4	46.1	50.3	46.2	37.4	45.2
	Avg. % Asian	10.5	2.3	2.3	1.5	1.0	0.9
	Avg. % Nat. Am.	0.4	0.5	0.4	0.5	0.5	0.5
	Avg. % Black	8.0	17.0	12.5	19.4	19.4	17.0
	Avg. % Pac. Isl.	0.2	0.1	0.1	0.1	0.1	0.3
	Avg. % Latinx	35.6	47.7	52.0	53.8	52.5	48.3
	Avg. % Multi-racial	2.4	1.8	1.5	1.4	1.2	2.1
	Avg. % White	42.9	30.5	31.1	23.3	25.3	30.8
Texas Charter Schools	Avg. % Econ. Dis.	70.3	69.1	70.4	62.8	81.2	59.0
	Avg. % LEP	5.8	10.9	6.8	8.6	7.6	6.3
	Avg. % SPED	2.6	9.9	5.9	11.5	93.1	12.6
	Avg. % Gifted	10.5	2.0	6.3	1.2	0.0	0.9
	Avg. % Female	55.5	52.9	52.8	50.5	39.1	43.8
	Avg. % Asian	8.2	1.3	2.7	1.3	4.3	0.8
	Avg. % Nat. Am.	0.1	0.6	0.3	0.4	0.0	0.7
	Avg. % Black	9.2	17.9	13.6	18.4	8.7	16.2
	Avg. % Pac. Isl.	0.0	0.1	0.1	0.2	0.0	0.4
	Avg. % Latinx	73.7	60.4	69.7	50.3	56.5	47.3
	Avg. % Multi-racial	0.6	1.1	0.8	1.9	0.0	1.5
	Avg. % White	8.1	18.7	12.9	27.5	30.4	33.2
All Texas Public Schools	Avg. % Econ. Dis.	33.2	60.5	55.1	68.9	73.2	62.9
	Avg. % LEP	1.5	7.9	5.7	10.1	7.9	5.2
	Avg. % SPED	1.0	15.5	5.2	17.8	91.3	18.6
	Avg. % Gifted	27.8	3.8	7.2	2.9	0.4	1.4
	Avg. % Female	51.6	46.3	50.4	46.9	37.4	44.8
	Avg. % Asian	10.4	2.3	2.3	1.4	1.1	0.9
	Avg. % Nat. Am.	0.4	0.5	0.4	0.5	0.5	0.6
	Avg. % Black	8.0	17.0	12.6	19.3	19.3	16.8
	Avg. % Pac. Isl.	0.2	0.1	0.1	0.1	0.1	0.3
	Avg. % Latinx	37.1	48.3	52.3	53.3	52.5	48.0
	Avg. % Multi-racial	2.3	1.7	1.5	1.5	1.2	1.9
	Avg. % White	41.5	30.0	30.8	23.9	25.3	31.5

Table 8. Coefficients from school-level multinomial logistic regression model using k-means clustering results as the outcome variable (Equation 18).

	Adv. Est (SE) Sig	Basic Est (SE) Sig	Trans. Est (SE) Sig	SPED Est (SE) Sig	Exit Est (SE) Sig
Charter	0.64 (0.21) **	0.71 (0.15) ***	1.20 (0.14) ***	-0.76 (0.41) .	0.98 (0.17) ***
Econ. Dis.	-1.31 (0.37) ***	-0.07 (0.26)	-0.10 (0.26)	0.80 (0.35) *	-0.04 (0.34)
SPED	-4.20 (0.12) ***	1.41 (0.40) ***	3.40 (0.28) ***	7.49 (0.33) ***	3.62 (0.34) ***
Gifted	1.29 (0.53) *	-2.22 (0.59) ***	-1.64 (0.63) **	0.03 (0.20)	-5.71 (0.17) ***
LEP	1.66 (0.57) **	1.57 (0.50) **	1.86 (0.43) ***	-0.80 (0.12) ***	1.62 (0.56) ***
MS Math	2.02 (0.32) ***	-0.53 (0.27) .	-1.63 (0.30) ***	1.65 (0.45) ***	-2.20 (0.46) ***
Asian	1.31 (0.64) *	1.60 (0.63) *	1.61 (0.69) *	-2.82 (0.15) ***	1.23 (0.17) ***
Black	-0.56 (0.49)	0.27 (0.37)	-0.81 (0.35) *	-7.36 (0.40) ***	-3.97 (0.46) ***
Latinx	-1.17 (0.39) **	-0.55 (0.28) .	-1.69 (0.28) ***	-7.17 (0.32) ***	-4.42 (0.37) ***
Mutli-race	3.42 (0.03) ***	0.90 (0.03) ***	1.68 (0.03) ***	-3.40 (0.01) ***	2.64 (0.01) ***
Nat. Am.	3.12 (0.01) ***	-0.69 (0.01) ***	-1.42 (0.01) ***	-0.44 (0.01) ***	-6.42 (0.01) ***
Pac. Is.	1.24 (0.01) ***	2.00 (0.01) ***	7.69 (0.01) ***	-10.2 (0.01) ***	3.54 (0.01) ***
White	-2.34 (0.26) ***	-0.07 (0.20)	-1.97 (0.21) ***	-6.66 (0.29) ***	-3.92 (0.28) ***
Schl. Size	.0007 (.0002) **	-.0003 (.0002)	-.002 (.0003) ***	.0005 (.0003)	-.002 (.0003) ***
# Comm	0.05 (0.06)	-0.05 (0.05)	0.36 (0.05) ***	0.84 (0.07) ***	0.85 (0.06) ***

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$

As with the other school-level models, the school-level demographic variables given in Table 8 (e.g., race/ethnicity, SPED, gifted, economic disadvantage, etc.) are a proportion of the total school population categorized as belonging to that demographic group and school size is measured in units of individual students. In Table 8, there is also a coefficient for the number of communities identified through the multilevel community detection algorithm applied to each school network.

Controlling for school level demographics, school size, and the number of communities identified within each Texas public school and relative to the odds of offering a "college preparatory" course sequence, charter schools are associated with an 86% increase in the likelihood of offering advanced course sequences, a 103% increase in the likelihood of offering a basic course sequences, a 232% increase in the likelihood of offering course sequences associated with transitions, and a 166% increase in the likelihood of offering course sequences associated with exit. These increases are all statistically significant. Charter schools are also associated with a 53% decrease in the likelihood of offering SPED tracks, but this is not statistically significant at the $p < 0.05$ level.

Table 9. Coefficients for student level hierarchical logistic regression models specified by Equation 20.

Log-Odds	**Adv. to CPrep**			**Basic to CPrep**			**Exit to CPrep**			**Trans. to CPrep**			**SPED to CPrep**		
	Est.	*SE*	*Sig.*	*Est.*	*SE*	*Sig.*	*Est.*	*SE*	*Sig.*	*Est.*	*SE*	*Sig.*	*Est.*	*SE*	*Sig.*
Intercept	-13.4	0.43	***	-3.44	0.20	***	-11.0	<.01	***	-4.85	0.26	***	-15.2	0.91	***
Charter	4.13	1.09	***	3.16	0.53	***	5.40	<.01	***	6.88	0.47	***	-0.03	0.84	
Econ. Dis.	-0.57	0.04	***	0.54	0.02	***	1.03	<.01	***	0.56	0.05	***	0.35	0.15	*
LEP	-0.86	0.08	***	0.73	0.03	***	0.36	<.01	***	0.52	0.05	***	0.03	0.20	
SPED	-1.08	0.10	***	2.21	0.03	***	2.60	<.01	***	2.29	0.03	***	9.38	0.46	***
Gifted	1.33	0.04	***	-0.46	0.04	***	-0.84	<.01	***	-0.77	0.12	***	-3.01	2.08	
Asian	1.18	0.20	***	-0.84	0.13	***	-0.31	<.01	***	-1.06	0.25	***	-3.06	1.02	**
Black	-0.51	0.20	*	0.04	0.12		-0.06	<.01	***	-0.42	0.20	*	-1.33	0.69	.
Pac. Isl.	-0.14	0.39		-0.86	0.29	**	-0.20	<.01	***	-0.71	0.50		1.72	1.93	
Latinx	-0.45	0.20	*	-0.01	0.12		-0.27	<.01	***	-0.62	0.20	**	-1.43	0.68	*
Multi-racial	-0.06	0.21		-0.20	0.13		0.90	<.01	***	-0.57	0.24	*	-1.53	0.80	.
White	-0.19	0.19		-0.22	0.12	.	0.48	<.01	***	-0.62	0.20	**	-1.41	0.68	*
MS Math	5.55	0.03	***	-0.90	0.03	***	0.19	<.01	***	-1.57	0.11	***	-7.80	2.50	**

Log-Odds	**Adv. to Basic**			**Exit to Basic**			**Trans. to Basic**			**SPED to Basic**			**Adv. to Trans.**		
	Est.	*SE*	*Sig.*	*Est.*	*SE*	*Sig.*	*Est.*	*SE*	*Sig.*	*Est.*	*SE*	*Sig.*	*Est.*	*SE*	*Sig.*
Intercept	-12.0	0.54	***	-8.56	1.49	***	-3.29	0.64	***	-16.5	0.99	***	-3.75	<.01	***
Charter	1.39	0.86		4.04	0.68	***	9.14	1.22	***	1.70	2.35		-5.81	<.01	***
Econ. Dis.	-1.05	0.06	***	-0.07	0.30		0.35	0.08	***	-0.23	0.15		-1.08	<.01	***
LEP	-0.67	0.11	***	-0.79	0.59		-0.01	0.11		-0.45	0.27	.	-0.82	<.01	***
SPED	-2.34	0.12	***	1.28	0.26	***	0.95	0.07	***	7.20	0.39	***	-2.55	0.13	***
Gifted	1.91	0.10	***	-0.24	0.86		-0.84	0.20	***	-11.7	7.93		1.60	<.01	***
Asian	0.88	0.29	**	0.18	1.56		-1.31	0.54	*	0.84	0.97		0.19	<.01	***
Black	-1.60	0.29	***	-0.35	1.41		-0.34	0.42		0.68	0.76		-1.90	<.01	***
Pac. Isl.	-3.54	0.59	***	2.17	1.86		-1.00	1.89		-27.0	12.3	*	-1.44	0.66	*
Latinx	-1.52	0.28	***	-0.33	1.39		-0.53	0.42		1.24	0.76		-1.53	<.01	***
Multi-racial	-1.02	0.32	**	-0.40	1.64		-0.26	0.48		1.51	0.92		-1.24	<.01	***
White	-0.91	0.28	**	-0.14	1.41		-0.44	0.42		0.94	0.76		-1.32	<.01	***
MS Math	6.10	0.07	***	-0.49	0.57		-1.81	0.15	***	-7.01	3.25	*	5.56	<.01	***

Table 9 (continued). Coefficients for student level hierarchical logistic regression models specified by Equation 20.

Log-Odds	**Exit to Trans.**			**SPED to Trans.**			**Exit to Adv.**			**SPED to Adv.**			**SPED To Exit.**		
	Est.	*SE*	*Sig.*	*Est.*	*SE*	*Sig.*	*Est.*	*SE*	*Sig.*	*Est.*	*SE*	*Sig.*	*Est.*	*SE*	*Sig.*
Intercept	-3.64	0.58	***	-7.68	1.04	***	-6.99	2.27	**	-5.28	1.31	***	0.28	0.78	
Charter	1.59	0.21	***	-3.02	0.86	***	7.19	1.44	***	-6.30	1.41	***	-1.33	0.18	***
Econ. Dis.	-0.84	0.23	***	-0.13	0.23		0.94	0.33	**	1.37	0.37	***	0.04	0.18	
LEP	-1.06	0.60	.	-0.59	0.31	.	-0.17	0.66		-0.70	0.64		2.63	0.97	**
SPED	1.80	0.21	***	6.17	0.39	***	4.81	0.44	***	11.1	0.59	***	1.74	0.20	***
Gifted	0.12	0.60		-1.11	1.91		-2.17	0.97	*	-3.68	1.74	*	-1.12	0.57	.
Asian	-0.17	0.66		-3.28	1.37	*	-1.52	2.40		-3.05	1.53	*	-0.80	0.93	
Black	-2.79	0.64	***	-3.21	0.87	***	0.64	2.21		-3.20	1.39	*	2.05	0.79	**
Pac. Isl.	-12.5	420		-0.24	2.40		2.16	2.79		-6.91	6.68		-18.7	3083	
Latinx	-2.29	0.58	***	-3.10	0.86	***	0.15	2.19		-3.45	1.34	*	2.01	0.78	**
Multi-racial	-4.82	2.75	.	-5.08	1.17	***	0.71	2.32		-3.14	1.72	.	4.18	1.75	*
White	-1.87	0.58	**	-3.06	0.87	***	0.78	2.19		-2.91	1.32	*	1.16	0.77	
MS Math	0.36	0.38		-2.62	1.80		-3.88	0.36	***	-8.87	0.95	***	-2.02	0.27	***

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$

The fifteen student-level models with coefficients provided in Table 9 compare the probabilities of students in Texas public schools following the course-sequences identified through community detection and defined through k-means clustering as a function of their demographic backgrounds and charter school enrollment. These models also control for campus level variation through random-effects coefficients for each campus included in this study.

At the student-level, enrollment in a charter school is associated with statistically significant increases in the likelihood that a student follows advanced, basic, transition, and exit course-sequences relative to the college preparatory course sequence. There is no statistically significant difference in the likelihood of a student enrolling in a SPED course sequence relative to a college preparatory sequence in charter schools. Relative to a basic course sequence, there is no statistically significant difference in the likelihood that a student enrolls in an advanced or SPED sequence, however, charter school students are statistically significantly more likely to enroll in exit and transition course sequences.

In analyzing the output from Table 9, the most likely course-sequences for charter school students are associated with exit and transition, while the least likely course-sequences for charter school students are the college preparatory and SPED course sequences.

Chapter 5: Discussion

As articulated in the introduction, the goals of this study are twofold: 1) to characterize differences in STEM course-taking options between charter and non-charter schools; and 2) to examine differences in student's STEM course-taking patterns in charter and non-charter secondary schools. Employing methods with principles rooted in the soft condensed matter community in physics allowed for common STEM course-sequences in Texas charter and non-charter schools to be identified, which could subsequently be analyzed using statistical methods. Differences in STEM course-offerings between charter and non-charter schools were investigated using school-level analyses, while differences in students' STEM course-taking patterns in Texas charter and non-charter schools were investigated using student-level analyses comparing the probabilities of students enrolling in different course sequences. A discussion of the primary findings related to the two research foci of this study are included in the two following subsections.

STEM Course-Offerings in Charter and Non-Charter Schools

A hierarchical model exploring the differences in the number of STEM course-sequences between charter and non-charter schools, as identified by the number of individual communities uncovered in each Texas public school, suggests there are no sector differences in the number of course-sequences offered when controlling for school level demographics and the size of the cohort population (Equation 17 and Table 9). Interestingly, increases in the percentage of gifted students in a given school and in the percentage of students who enrolled in high-school math in 8^{th} grade in a school are associated with decreases in the number of course-sequences offered in that school.

Although the number of individual STEM course-sequences was not found to differ between charter and non-charter schools in Texas, there are sector differences in the *kinds* of STEM course-sequences offered. After constructing school-level sociograms in which schools are connected by the number of STEM courses common to each pair of schools, three communities of schools are identified with their associated courses listed in Table 3.

Relative to the community with the widest range of STEM courses offered (including advanced and SPED courses), charter schools are less likely to offer STEM courses heavily tailored for SPED students. There is no statistically significant difference in a charter school belonging to a community with a wide-range of STEM course-options and limited, non-SPED STEM courses, however.

In contrasts to narratives that promote charter schools as educational institutions capable of offering novel instruction and curricula, results from the hierarchical model and analysis of the school-level community detection suggest that charter schools and non-charter schools are more alike in course-offerings than they are different. Charter schools are not more or less likely than non-charter schools to offer STEM courses that are minimal, consisting of only staple courses with few electives, nor are they more or less likely to offer expansive course offerings, which include advanced STEM courses and STEM courses tailored for SPED students. An important exception is that charter schools are less likely than non-charter schools to offer STEM courses that are heavily catered toward SPED students.

In statistically analyzing results from student-level community detection, in which k-means clustering is used to group STEM course-sequences with common attributes, the course-sequences in charter schools are more likely to be “advanced” and “basic” when compared to course-sequences that are considered college preparatory. In addition, STEM course sequences in charter schools are more often associated with mobility (transition and exit) than course sequences in non-charter schools.

As studies exploring differences in student achievement have noted, these differences appear to be highly contextual. While the finding that charter schools are simultaneously more likely to offer “advanced” and “basic” course sequences relative to college preparatory course sequences may seem at first counterintuitive, it is possible that the charter schools offering “advanced” course sequences are contextually different than charter schools offering “basic” tracks. This is a speculation that warrants further consideration and may lend additional insight into the different kinds of academic programs offered in various charter schools. Specifically, it may be that charter schools target different populations: college preparatory charter schools may target populations who are pushed into advanced STEM coursework; whereas other schools—perhaps charter schools serving high populations of students deemed "at-risk" for dropping out of—may offer basic and minimal STEM coursework. By contrast, non-charter schools do not have the freedom to recruit specific subsets of students and thus do not tailor academic programming to meet the needs of a specific subset of students. As such, course-offerings in non-charter schools are broader than those in charter schools, as they cater to a more heterogenous student population.

STEM Course-Taking Patterns in Charter and Non-Charter Schools

STEM course-offering differences between charter and non-charter schools as identified during school-level analyses align with differences in student-level STEM course-taking pattern differences between charter and non-charter schools in Texas. This is a sensible finding, as course-offerings within a given school necessarily constrain student course-taking options. Results from hierarchical logistic regression models suggest that students in charter schools are more likely enrolled in course sequences characterized by high mobility (e.g., transfer and dropping out) than are students in non-charter schools. In addition, students in charter schools more often take course sequences that have a greater number of advanced STEM courses or that are characterized by minimal STEM coursework. Relative to these four sets of courses, students in charter schools are less likely to take "college preparatory" STEM course-sequences or STEM course-sequences that cater to SPED populations.

These results are consistent with other studies finding that students in charter schools are more-likely than students in non-charter schools to be enrolled in advanced course sequences (Berends & Donaldson, 2016). However, the data analyzed in this study compare specific course-taking patterns in all Texas public schools, whereas Berends and Donaldson (2016) infer ability groups by analyzing survey responses from teachers in which these teachers describe their instructional practice. An important way in which this work expands upon prior work investigating ability grouping is that this work offers insight into the specific STEM courses students take in charter and non-charter schools. It is likely that some of the differences observed may vary by each school, as different charter schools serve different populations and may very well have different course-offerings to meet the needs of these students.

As articulated in the preceding section, these seemingly contradictory findings may reflect school-level differences within charter schools. The specific academic programs in some charter schools may be tailored such that students take more advanced courses, while the academic programs in other schools may serve to give students only basic preparation in STEM. In addition, charter school cohorts are smaller than cohorts in non-charter schools. As such, STEM course-taking patterns in non-charter schools cater to a broader group of students, whereas the course-taking patterns in charter schools, while more limited, are designed to meet the specific needs of the students they serve.

Chapter 6: Conclusion

This study employed social network analysis, the multilevel community detection algorithm, and inferential statistical models in order to explore STEM course-offerings and student STEM course-taking patterns in Texas charter and non-charter secondary schools. Community detection allowed for students sharing several courses in Texas public schools to be identified within each school included in this study. In analyzing the specific course-taking records of students associated with identified communities, the k-means clustering algorithm allowed for the categorization of these communities into six groups. Results indicate that in general, charter schools and non-charter schools offer similar STEM courses to students, except that charter schools are less likely to have course-offerings tailored for SPED students. There are sector differences in the associated course-sequences in charter and non-charter schools. Of the six groups of course-sequences identified in this study, charter schools are most likely to offer course-sequences associated with mobility (transfer and dropping out). By contrast, charter schools are least likely to offer course-sequences that have been defined herein as college preparatory (meaning students take several STEM courses, but not advanced electives) and course-sequences tailored for SPED students. Relative to the "college preparatory" course sequence, charter schools are more likely to offer tracks characterized by a high number of advanced coursework in STEM or characterized by minimal course-taking in STEM.

The methods employed within this thesis offer a promising and novel way to identify and analyze differences in course-offerings and course-taking patterns in Texas charter and non-charter schools. While these differences were not analyzed with respect to differences in student outcomes—either test score increases, college enrollment, or labor market outcomes—characterization of programmatic differences between charter and non-charter schools can lend insight into future studies that specifically explore these differential student outcomes. Moreover, the results of this study suggest that while general course-offerings between charter and non-charter schools may not embody the "innovation" promised by school choice advocates, results suggest there are stark differences between charter and non-charter schools in student STEM course-taking patterns. Investigating these differences more deeply is an important next step for research on charter schools.

LIMITATIONS AND FUTURE DIRECTIONS

An area for future research is to explore community detection in heavily weighted networks. Many studies referenced in this thesis analyzed the efficacy of community detection algorithms using sparse networks that are unweighted and undirected. The data analyzed in this work were characterized by dense networks with heterogeneously distributed weights, and the efficacy of community detection algorithms in dealing with these networks has not been systematically explored to the same degree they have been explored on undirected, unweighted, and sparse networks. While the community detection algorithms explored in this study offer promising methodologies for exploring and analyzing educational data in ways that are not often done in educational research, there is still a need for researchers to investigate the efficacy of community detection algorithms to identify communities within the dense, weighted networks characteristic of empirical data.

Moreover, while this work sheds insight into the different types of course-taking sequences that are typical of charter school students, it does not explore how students within individual charter and non-charter schools are grouped according to ability. This work is a necessary first step in identifying general differences in course offerings and student-course taking patterns between charter and non-charter schools, however, future work should investigate how students within these schools are grouped in order to see if ability grouping practices differ by school sector.

Bibliography

Anderson, M. (2018, November 15). Charter Schools Program. Retrieved from https://innovation.ed.gov/files/2015/03/CSP-DCL.pdf

Barrat, A., Barthélemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences*, *101*(11), 3747–3752. https://doi.org/10.1073/pnas.0400087101

Begg, C. B., & Gray, R. (1984). Calculation of Polychotomous Logistic Regression Parameters Using Individualized Regressions. *Biometrika*, *71*(1), 11–18. https://doi.org/10.2307/2336391

Bendinelli, A. J., & Marder, M. (2012). Visualization of longitudinal student data. *Physical Review Special Topics - Physics Education Research*, *8*(2), 020119. https://doi.org/10.1103/PhysRevSTPER.8.020119

Berends, M. (2015). Sociology and School Choice: What We Know After Two Decades of Charter Schools. *Annual Review of Sociology*, *41*(1), 159–180. https://doi.org/10.1146/annurev-soc-073014-112340

Berends, M., & Donaldson, K. (2016). Does the Organization of Instruction Differ in Charter Schools? Ability Grouping and Students' Mathematics Gains. *Teachers College Record*, *118*(11), 1.

Berends, M., Goldring, E., Stein, M., & Cravens, X. (2010). Instructional conditions in charter schools and students' mathematics achievement gains. *American Journal of Education*, *116*(3), 303–335.

Bierlein, L. A., & Mulholland, L. A. (1994). The promise of charter schools. *Educational Leadership*, *52*, 34.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Borgatti, S. P., & Ofem, B. (2010). Overview: Social Network Analysis and Theory. In A. J. Daly & J. W. Little (Eds.), *Social Network Theory and Educational Change*. Cambridge, Mass: Harvard Education Press.

Bruna, J., & Li, X. (2017). Community Detection with Graph Neural Networks. *ArXiv:1705.08415 [Stat]*. Retrieved from http://arxiv.org/abs/1705.08415

Carnevale, A. P., Cheah, B., & Hanson, A. R. (2015). *The Economic Value of College Majors*. Georgetown University Center on Education and the Workforce.

Carolan, B. (2014). *Social Network Analysis and Education: Theory, Methods & Applications*. 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc. https://doi.org/10.4135/9781452270104

Charter Schools. (n.d.). Retrieved April 30, 2018, from https://innovation.ed.gov/what-we-do/charter-schools/

Clark, M. A., Gleason, P. M., Tuttle, C. C., & Silverberg, M. K. (2015). Do Charter Schools Improve Student Achievement? *Educational Evaluation and Policy Analysis*, *37*(4), 419–436. https://doi.org/10.3102/0162373714558292

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, *70*(6). https://doi.org/10.1103/PhysRevE.70.066111

Csardi, G. (2015). igraph: Network Analysis and Visualization (Version 1) [R]. Retrieved from http://igraph.org

Curto, V. E., & Fryer, R. G. (2014). The Potential of Urban Boarding Schools for the Poor: Evidence from SEED. *Journal of Labor Economics*, *32*(1), 65–93. https://doi.org/10.1086/671798

Danon, L., Díaz-Guilera, A., Duch, J., & Arenas, A. (2005). Comparing community structure identification. *Journal of Statistical Mechanics: Theory and Experiment*, *2005*(09), P09008. https://doi.org/10.1088/1742-5468/2005/09/P09008

DeVos, B. Secretary’s Proposed Supplemental Priorities and Definitions for Discretionary Grant Programs, 82 FR 47484 § (2017).

Dobbie, W. S., & Fryer, J. R. G. (2016). *Charter Schools and Labor Market Outcomes* (Working Paper No. 22502). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w22502

Doyle, M. C., & Feldman, J. (2006). Student Voice and School Choice in the Boston Pilot High Schools. *Educational Policy*, *20*(2), 367–398. https://doi.org/10.1177/0895904805284051

Feder, M. (2012, December 18). One Decade, One Million more STEM Graduates. Retrieved May 12, 2016, from https://www.whitehouse.gov/blog/2012/12/18/one-decade-one-million-more-stem-graduates

Finn, A. S., Kraft, M. A., West, M. R., Leonard, J. A., Bish, C. E., Martin, R. E., … Gabrieli, J. D. E. (2014). Cognitive Skills, Student Achievement Tests, and Schools. *Psychological Science*, *25*(3), 736–744. https://doi.org/10.1177/0956797613516008

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, *486*(3), 75–174. https://doi.org/10.1016/j.physrep.2009.11.002

Friedkin, N. E., & Thomas, S. L. (1997). Social Positions in Schooling. *Sociology of Education*, *70*(4), 239–255. https://doi.org/10.2307/2673266

Friedman, M. (2009). *Capitalism and Freedom: Fortieth Anniversary Edition*. Chicago, IL: University of Chicago Press. Retrieved from http://ebookcentral.proquest.com/lib/utxa/detail.action?docID=432222

Fruchterman, T. M. J., & Reingold, E. M. (1990). *Graph drawing by force-directed placement*. Urbana, IL: Dept. of Computer Science, University of Illinois at Urbana-Champaign.

Girvan, M., & Newman, M. E. J. (2002). Community Structure in Social and Biological Networks. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(12), 7821–7826.

Gleason, P., Clark, M., Tuttle, C., & Dwoyer, E. (2010). *The evaluation of charter school impacts*. Washington, DC: U.S. Dept. of Education, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences.

Guggenheim, D. (2010). *Waiting for "Superman."* Hollywood, CA: Paramount Home Entertainment (Firm).

Guthrie, M. W. (2018). *Grouping and comparing Texas high schools through machine learning and visualization techniques* (Dissertation). The University of Texas at Austin, Austin, Texas.

Heck, R. H., Price, C. L., & Thomas, S. L. (2004). Tracks as Emergent Structures: A Network Analysis of Student Differentiation in a High School. *American Journal of Education*, *110*(4), 321–353. https://doi.org/10.1086/422789

Henig, J. R. (1995). *Rethinking School Choice, Limits of the Market Metaphor* (With a New afterword by the author). Berlin, Boston: Princeton University Press. https://doi.org/10.1515/9781400821037

Huerta, L. A., & Zuckerman, A. (2009). An Institutional Theory Analysis of Charter Schools: Addressing Institutional Challenges to Scale. *Peabody Journal of Education*, *84*(3), 414–431.

Jabbar, H. (2015). "Every Kid Is Money": Market-Like Competition and School Leader Strategies in New Orleans. *Educational Evaluation and Policy Analysis*, *37*(4), 638–659. https://doi.org/10.3102/0162373715577447

Jabbar, H. (2016). The Visible Hand: Markets, Politics, and Regulation in Post-Katrina New Orleans. *Harvard Educational Review*, *86*(1), 1–26.

Katz, N., Lazer, D., Arrow, H., & Contractor, N. (2004). Network Theory and Small Groups. *Small Group Research*, *35*(3), 307–332. https://doi.org/10.1177/1046496404264941

Lacireno-Paquet, N., Holyoke, T. T., Moser, M., & Henig, J. R. (2002). Creaming Versus Cropping: Charter School Enrollment Practices in Response to Market Incentives. *Educational Evaluation and Policy Analysis*, *24*(2), 145–158. https://doi.org/10.3102/01623737024002145

Lancichinetti, A., & Fortunato, S. (2009). Benchmarks for testing community detection algorithms on directed and weighted graphs with overlapping communities. *Physical Review E*, *80*(1). https://doi.org/10.1103/PhysRevE.80.016118

Lancichinetti, A., Fortunato, S., & Radicchi, F. (2008). Benchmark graphs for testing community detection algorithms. *Physical Review E*, *78*(4), 046110. https://doi.org/10.1103/PhysRevE.78.046110

Lancichinetti, A., Radicchi, F., Ramasco, J. J., & Fortunato, S. (2011). Finding Statistically Significant Communities in Networks. *PLoS ONE*, *6*(4). https://doi.org/10.1371/journal.pone.0018961

Lubienski, C. (2003). Innovation in Education Markets: Theory and Evidence on the Impact of Competition and Choice in Charter Schools. *American Educational Research Journal*, *40*(2), 395–443. https://doi.org/10.3102/00028312040002395

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. Presented at the Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics, The Regents of the University of California. Retrieved from https://projecteuclid.org/euclid.bsmsp/1200512992

Marder, M., & Bansal, D. (2009). Flow and diffusion of high-stakes test scores. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(41), 17267–17270. https://doi.org/10.1073/pnas.0812221106

McDermott, K. A., & Nygreen, K. (2013). Educational New Paternalism: Human Capital, Cultural Capital, and the Politics of Equal Opportunity. *Peabody Journal of Education*, *88*(1), 84–97. https://doi.org/10.1080/0161956X.2013.752634

McFarland, D. A. (2006). Curricular Flows: Trajectories, Turning Points, and Assignment Criteria in High School Math Careers. *Sociology of Education*, *79*(3), 177–205. https://doi.org/10.1177/003804070607900301

Metcalf, H. (2010). Stuck in the Pipeline: A Critical Review of STEM Workforce Literature. *InterActions: UCLA Journal of Education and Information Studies*, *6*(2). Retrieved from http://escholarship.org/uc/item/6zf09176

Modica, M. (2015). “My Skin Color Stops Me from Leading”: Tracking, Identity, and Student Dynamics in a Racially Mixed School. *International Journal of Multicultural Education*, *17*(3), 76–90.

Mungal, A. S. (2016). Teach for America, Relay Graduate School, and the Charter School Networks: The Making of a Parallel Education Structure. *Education Policy Analysis Archives*, *24*(12–18), 1–26. https://doi.org/10.14507/epaa.24.2037

NAPCS. (2018). Charter School Data Dashboard. Retrieved October 12, 2016, from http://dashboard.publiccharters.org/dashboard/schools/year/2014

National Center for Education Statistics. (2018). *Digest of Education Statistics, 2016* (No. NCES 2017-094). Washington, D.C.: U.S. Department of Education, Institute of Education Sciences. Retrieved from http://nces.ed.gov/programs/digest/d15/tables/dt15_322.30.asp

National Science Board. (2016). Science and Engineering Indicators 2016. Retrieved October 24, 2016, from https://www.nsf.gov/statistics/2016/nsb20161/#/digest

Newman, M. E. J. (2004). Fast algorithm for detecting community structure in networks. *Physical Review E*, *69*(6), 066133. https://doi.org/10.1103/PhysRevE.69.066133

Newman, M. E. J. (2006). Modularity and Community Structure in Networks. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(23), 8577–8582.

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, *69*(2). https://doi.org/10.1103/PhysRevE.69.026113

Office of the Press Secretary. (2009, November 23). President Obama Launches "Educate to Innovate" Campaign for Excellence in Science, Technology, Engineering & Math (Stem) Education. Retrieved February 9, 2016, from https://www.whitehouse.gov/the-press-office/president-obama-launches-educate-innovate-campaign-excellence-science-technology-en

Reichardt, J., & Bornholdt, S. (2006). Statistical mechanics of community detection. *Physical Review E*, *74*(1), 016110. https://doi.org/10.1103/PhysRevE.74.016110

Rothwell, J. (2013). *The Hidden STEM Economy*. Brookings Institution.

Salzman, H. (2013). What Shortages? The Real Evidence About the STEM Workforce. *Issues in Science & Technology*, *29*(4), 58–67.

Schweitzer, F. (2018). Sociophysics. *Physics Today*, *71*(2), 40–46. https://doi.org/10.1063/PT.3.3845

Teitelbaum, M. S. (2014, March 19). The Myth of the Science and Engineering Shortage. *The Atlantic*. Retrieved from http://www.theatlantic.com/education/archive/2014/03/the-myth-of-the-science-and-engineering-shortage/284359/

Toma, E., & Zimmer, R. (2012). Two decades of charter schools: Expectations, reality, and the future. *Economics of Education Review*, *31*(2), 209–212. https://doi.org/10.1016/j.econedurev.2011.10.001

Tuttle, C. C., Gleason, P., & Clark, M. (2012). Using lotteries to evaluate schools of choice: Evidence from a national study of charter schools. *Economics of Education Review*, *31*(2), 237–253. https://doi.org/10.1016/j.econedurev.2011.07.002

West, M. R., Gabrieli, C. F. O., Finn, A. S., Kraft, M. A., & Gabrieli, J. D. E. (2014). What Effective Schools Do. *Education Next*, *14*(4), 72–79.

Winters, M. A. (2012). Measuring the effect of charter schools on public school student achievement in an urban environment: Evidence from New York City. *Economics of Education Review*, *31*(2), 293–301. https://doi.org/10.1016/j.econedurev.2011.08.014

Winters, M. A. (2015). Understanding the Gap in Special Education Enrollments between Charter and Traditional Public Schools: Evidence from Denver, Colorado. *Educational Researcher*, *44*(4), 228–236.

Wu, B. (2014). The weighted version of the handshaking lemma with an application. *Journal of Inequalities and Applications*, *2014*(1), 1–5. https://doi.org/10.1186/1029-242X-2014-351

Yang, Z., Algesheimer, R., & Tessone, C. J. (2016). A Comparative Analysis of Community Detection Algorithms on Artificial Networks. *Scientific Reports*, *6*, 30750. https://doi.org/10.1038/srep30750

Zimmer, R., Gill, B., Booker, K., Lavertu, S., & Witte, J. (2012). Examining charter student achievement effects across seven states. *Economics of Education Review*, *31*(2), 213–224. https://doi.org/10.1016/j.econedurev.2011.05.005